PROJECT ON ANALYSIS OF COMMUNITY
LEVEL NUTRITION PROGRAMS

Volume I:

Final Report: Analysis of Community-Level Nutrition Programs

William D. Drake, Ph.D.
Roy I. Miller, Ph.D.
Margaret Humphrey, M.S.W.

October 1980

Office of Nutrition

U.S. Agency for International Development
Washington, D.C.

Community Systems Foundation
1130 Hill Street
Ann Arbor, Michigan 48104

3783

This volume is one of a series of eight produced by the Project on Analysis of Community-Level Nutrition Programs. The four year research project utilized an interdisciplinary team from many parts of the world.

Principal Staff:

William D. Drake, Ph.D. Research Project Director. Professor of Resource Planning & Conservation, University of. Michigan, and Associate, Community Systems Foundation.

Roy I. Miller, Ph.D. Associate, Community Systems Foundation.

Margaret H. Humphrey, M.S.W. Associate, Community Systems Foundation.

John Nystuen, Ph.D. Professor of Geography, University of Michigan.

Donald A. Schön, Ph.D. Ford Professor of Urban Studies & Education, Massachusetts Institute of Technology.

Luis J. Fajardo, M.D. Professor of Medicine, Universidad del Valle, Cali, Colombia.

John O. Field, Ph.D. Associate Professor of Political Science and Nutrition, Tufts University.

Contributing Staff:

Within the Community Systems Foundation, many staff contributed generously to the project. They include: Barton Burkhalter, Dean Wilson, David Sahn, Robert Pestronk, Naomi Gottlieb, Martha Gregg, James Eckroad, David Nelson, Charles Eisendrath, Stephen Hynson, Than-Long Ton-That, Glen Melnick, Mutombo Mpanya, Barry Rose, Robert Tilden, Victor Santiago, Eileen Trzcinski, Janet Wilson and Gloria Quintero.

In addition to the Foundation staff, many individuals in the various field programs and in other institutions contributed to the effort. Their names and the role they played are cited in the appropriate volumes.

Project Reports:

Volume I Final Report: Analysis of Community-Level Nutrition Programs

Volume II Esperança: A Program to Improve Health and Nutrition In the Central Amazon

Volume III Kottar: Malnutrition, Intervention and Development in a South Indian District

Volume IV The Promotora Program in Candelaria
Part I: A Colombian Attempt to Control Malnutrition and Disease, 1968-1974
Part II: A Revisitation Two Years After Program End

Volume V The PRIMOPS Experience: Information Processing in the Design and Performance of a Health Care System

Volume VI Selected Analyses: Dominican Republic, Honduras, Indonesia, Thailand

Volume VII The Rapid Information Feedback System: A Mechanism for Promoting Community Learning

Volume VIII Internal Reports: Code Books, Growth Standards, Intake Standards and Other Documentation

Community Health Cell
Library and Documentation Unit
BANGALORE

PROJECT ON ANALYSIS OF COMMUNITY
LEVEL NUTRITION PROGRAMS

Volume I:

Final Report: Analysis of Community-Level Nutrition Programs

William D. Drake, Ph.D.
Roy I. Miller, Ph.D.
Margaret Humphrey, M.S.W.

October 1980

Office of Nutrition

U.S. Agency for International Development
Washington, D.C.

Community Systems Foundation
1130 Hill Street
Ann Arbor, Michigan 48104

COMMUNITY HEALTH CELL
LIBRARY
AND
DOCUMENTATION
UNIT
BANGALORE.
3783
NUT100

PREFACE

Attributing authorship of a report that includes the thoughts and perspectives of a great many people is a difficult task. The primary responsibility for the document rests on the three people whose major commitment over the past four years has been to this project: William D. Drake, principal investigator; Roy I. Miller; and Margaret H. Humphrey. However, the contributions of the many other people, who in some way touched this project, were essential for its success.

Almost everyone at Community Systems Foundation contributed time, and more importantly, wisdom to this project. Heading the list is John Nystuen, who played two critical roles—primary liaison with the PRIMOPS intervention, and constant participant in the conceptual development of the project. The clear thinking of Don Schon assisted in the development of a framework for organizing the entire effort. John Field added the perspective of the political scientist and made the KOTTAR data set and connection "happen." Dean Wilson provided a constant reminder of the need to blend empirical work with well-thought-out theory. Barton Burkhalter had a wealth of experience in evaluation work—especially in the area of nutrition. Luis Fajardo brought his knowledge of the medical reality of malnourishment and helped place analyses of data in the context of that reality.

Richard Paullin, Martha Gregg, Robert Pestronk, and David Sahn, the CSF people whose primary effort is directed toward publishing the <u>Nutrition Planning</u> journal, shared their knowledge of the literature. Similarly, other CSF regulars, James C. Eckroad and David Nelson, contributed the knowledge gained through their own activities. Even the expertise of the CSF Board of Trustees was sought and supplied as Frederick Goodman, Merrill Flood, Robert Farrell, and William J. Horvath addressed analytic issues of particular interest in their own work. Other

consultants to CSF were "tapped" when possible—Bill Bertrand, Alberto Pradilla, Gloria Quintero, and James Koopman.

The project also utilized some expert "student help" from the University of Michigan. Tasks ranging from data cleaning to analysis, literature review, and map making were tackled by Stephen Hynson, Than-Long Ton-That, Glenn Melnick, Mutombo Mpanya, Barry Rose, Robert Tilden, Victor Santiago, Eileen Trzcinski, and Janet Wilson.

Of course, nothing could have happened without the change agents—the field people who contributed their data, time, and experience. These include Jamie Rodriguez (Colombia), Alfredo Aguirre (Colombia), Ramiro Munoz (Colombia), Giovanni Acciari (Colombia), Fred Hartman (Brazil), Mary Hartman (Brazil), Steve Alexander (Brazil), Norma Pereira (Brazil), Kate Burns (Brazil), Chuck Post (Brazil), Father James Tombeur (India), Sister Godelieve VandeWalle (India), Zafulla Choudry (Bangladesh), and Thad Jackson (Bangladesh).

Phyllis Dobyns of Save The Children was the instrumental link to the several SAVE interventions reviewed during this project. SAVE field people included Chris and Gary Schaye (Dominican Republic), Jose Andy Rubi (Honduras), Alex Corpeno (Honduras), George Moore (Honduras), Marty and Ruchire Poland (Indonesia), Nancy Terreri (Indonesia), Siti Sundari (Indonesia), and Alton Straughan (Bangladesh).

Several of the data sets came to us from American colleagues who had developed those data sets in their own work. Stanley Gershoff and Robert McGandy conducted the rice fortification program in Thailand. Merlyn Vemury made the CARE data set available to us.

Finally, the US/AID staff in Washington was very helpful. Martin J. Forman, Director of the Office of Nutrition, spent many days with us sharing his knowledge of the field as well as his perspective as a US/AID official. Other AID staff, especially Carol Adelman, Harold Rice, Jim Levinson, Robert Pratt, Mary Ann Anderson, and Hope Sukin always had time to talk and a willingness to share their substantive knowledge as well as their administrative skills.

SUMMARY
ANALYSIS OF COMMUNITY-LEVEL NUTRITION PROGRAMS

I. INTRODUCTION

For the past four years, an interdisciplinary research team has been studying community-level nutrition interventions conducted in the developing world. The purpose of this project was to gain a clearer understanding of which community-level program characteristics result in improved nutritional status of children. A complementary objective has been to further develop the methodology of program evaluation, and report these findings in a manner which would assist both ongoing and future programs in accomplishing their goals.

The investigation has included a substantial review of the methodological literature, studies of existing project reports, numerous visits to ongoing projects and, most important, an in-depth analysis of seven different projects. All of these in-depth studies had substantial amounts of longitudinal anthropometric measures of child nutritional status, varying amounts of other health data, and socio-demographic, environmental, and program information. In several instances, the research team mounted a supplementary data-gathering effort to fill critical gaps in knowledge about the project and the setting in which it took place. In half of the cases, field teams were placed on site for several months at a time.

The numerical data from the projects were put through exhaustive error detection and correction procedures, both manual and computer based, and then placed in a newly developed data base management system. This system allowed the research team to retrieve efficiently subsets of the data for subsequent manipulation. Over 50 different anthropometric standards were compiled, and several of the most frequently utilized ones were entered into the data base management system, thereby facilitating

comparative analysis. The development of this approach to handling data was especially important because of the complexity of some of the required manipulations and the large number of variable observations—over four million. A description of this versatile data base management system and the many computer subroutines associated with error detection and correction is contained in Volume 7 of this report.

Numerous methodological techniques were used in the analysis of the data, ranging from very simple techniques appropriate on site in the field to the most sophisticated nonlinear multivariate procedures. In many instances, parallel analyses were performed on the same data, using different techniques to see the similarities and the differences in outcome. Efforts were also undertaken with varying levels of "cleanliness" in the data to see the relative merit of expending resources on the cleaning process. Members of the research team with different disciplinary backgrounds and viewpoints performed different styles of analysis in addition to alternative methodological techniques. These results were then reviewed with intervention project personnel and field staff as well as with the rest of the research team.

Considerable emphasis was placed on developing a comprehensive system for rapidly retrieving data from the field, converting it to machine-readable form, performing error detection and correction schemes, entering it into the data base management system, performing alternative analysis, and writing a brief report on the results. This system allowed us to experiment with various information feedback procedures. In the early stages of the research project, it took several months to carry out all of the aforementioned steps, but now, after three years of development, we have experienced turnaround times of less than one week for a typical set of data consisting of approximately two thousand observations.

The conclusions stemming from this research project fall into two general categories: (1) methodological; and (2) substantive. While it is useful to distinguish between the two types of conclusions, there are also areas of heavy overlap. That is, effective evaluative methodologies are dependent on the nature of the intervention, while the efficacy of the

intervention is, in part, determined by the type of self-evaluation and feedback utilized.

II. FINDINGS REGARDING EVALUATION METHODOLOGIES

Determining whether a community-level nutrition project helped reduce malnutrition is not a simple task. One must first measure the change in nutritional status in the community or participants. Then, one must attribute the observed change to the project. Neither task is simple.

A. Measuring Outcome

For convenience, let us partition our findings with regard to measuring change or outcome into three groups. The first group of findings identify problems inherent in the mathematics and definitions of change in nutritional status. These cannot be "corrected" but can, at least, be acknowledged.

1. In research strategy that compares groups—for example, participants to nonparticipants, or "befores" to "afters"--misclassification of cases will always create a situation where the observed difference between groups is less than the true or actual difference. When the usual procedures are used to define malnourishment by comparing observed anthropometrics to accepted reference standards, there is, by definition, some such misclassification. Some small but healthy children will be called malnourished; other large but sick children will be called normal.

2. The magnitude of the understatement of differences is directly related to the degree to which children are misclassified. Unfortunately, additional misclassification arises because measurement techniques are also imperfect. Therefore, inaccuracies in weighing children, determining their height, or calculating their age contribute further to the understatement of differences.

3. Some program impacts often go undetected because anthropometric measures of nutritional status fail to capture the complete effect of an intervention upon individuals—especially individuals who have suffered chronic malnourishment. Children severely malnourished for sustained periods in early life may simply be unable to achieve the recuperative growth necessary to display

nutritional improvement with regard to a growth standard. Thus, lack of growth does not directly indicate lack of program impact. Children suffering chronic malnutrition may experience improvements with regard to metabolic rate or level of activity but not with regard to growth. Consequently, analysis based on anthropometrics tends to **underestimate** the true effect of an intervention.

The second group of findings is directed toward the vagaries of statistical analysis. These illustrate how the subjective judgments of any analyst can bias the determination of outcome.

4. Results of analysis are influenced by the anthropometric standard applied to determine nutritional status. Local standards that account for genetic or cultural differences in populations may yield different nutritional profiles than would a generalized standard calibrated to healthy North American populations. The application of a sex-differentiated standard, rather than a unisex standard, can actually shift results from significant to insignificant.

5. Differences in nutritional profile resulting from different reference standards are further exacerbated by the selection of "cut-points" to define malnutrition. For many nutrition interventions, changes in the mean percent of standard achieved by the target population are only one of a series of important indicators of impact. Changes in the seriously malnourished (Grade III) or the children most at risk (Grades II and III) may be more relevant, and, of course, the definition of these groups depends on the choice of "cut-points."

6. Analytic results can be changed not only by changing reference standards and cut-points but also by changing the statistical tests and/or the "artistic" selection of variables to be included in the analysis. The rationale for the selection of tests and variables should be well stated by any analyst and, where possible, the model and associated statistical tests used in analysis should be stated in advance.

The last group of findings with regard to measuring outcome identify problems which are, in theory, correctable by proper field protocol. Realistically, however, one must acknowledge that in most interventions in the developing world, perfect field protocol will never be attained.

7. Most anthropometric assessments derive nutritional status by comparing an observed measurement with a reference standard specified by age of child. Although this assessment requires a precise determination of child-age in months, in many cultures of the developing world there are enormous problems associated with obtaining correct birthdates. Births frequently go unregistered because children are born at home, and incentives often exist to under-report age so that the child may remain eligible for program benefits.

8. Accurate measures of weight and height are also difficult to obtain in communuity-level interventions. Not only are the delicate scales used for weighing difficult to maintain in a field setting, but measures of height often become approximations as field workers struggle with squirming children.

Since it is unrealistic to presume the elimination in the field of these measurement problems, the analyst must recognize that underreporting of program effects will occur; and the greater the accuracy problems, the greater will be the underreporting as described in #1.

B. Attributing the Outcome to a Cause

The difficulties of determining whether there was an improvement in nutritional status of the target population is only one component of program evaluation. A more important issue is the attribution of the observed change to a cause: whether the cause is the intervention itself or some other phenomenon that happens to be occurring during the same time period. In all cases there are other possible reasons (besides the intervention itself) that compete with the program for explaining why the change occurs. While the list of all potential competing explanations is impossible to enumerate, seven general categories of competing explanations have been identified:

1. Changes in environment. The environment (social, economic, political, and physical) upon which each intervention is superimposed is constantly changing--sometimes gradually (e.g., inflation, introduction of other government services, etc.) and sometimes suddenly (e.g., drought, military upheaval, etc.). Changes in the environment often affect nutritional status more than any

intervention ever could.

2. Changes in population age. The demographic make-up of study populations changes over time. This is particularly important with respect to the aging of children. One obvious relationship randomly found in the data is the "cycle" most children live through of deteriorating nutritional status after weaning, followed by natural improvement. The net result of this cycle is that any group of two-year-olds will look "better" by the time they are four. Mortality, especially among those suffering from severe malnutrition, further complicates efforts to observe the same group of children at different points in time.

3. Change in target populations. Above and beyond the natural aging of populations, the composition of a target population changes—through in-migration, out-migration, and death.

4. Changes in selection criteria. In most programs, especially ones emphasizing service rather than research, participants are not selected randomly. Rather, they are selected because of willingness to participate, proximity to the service provider, religious affiliation, or caste. As the intervention proceeds, the selection criteria may change, with a consequent change in the apparent outcome.

5. Age of project. The skill level of staff members, the enthusiasm of staff members, the resources available to a program—these elements are constantly changing and could underlie observed changes in impact measures.

6. Changes in data reliability. Data-gathering procedures atrophy (measurement and recording gets sloppy) if data are not used. When data are used for decision-making, the reliability often improves. Either way, outcomes observed over time can reflect qualitative changes in the data and not true program-related changes.

7. Hawthorne effects. The mere presence of a program may cause changes in the behavior of a population, independent of the content of the program; these changes, in turn, may be sufficient to generate changes in the selected outcome measure.

The traditional scientific "solution" to resolving which competing explanation is the cause of program impact has been to use controlled experiments—either statistical controls or paired populations. However, our

investigations both of the literature and of our own data sets reveal no instance in the developing world where controlled experiments have been successful in resolving—unambiguously—the attribution of an outcome to a cause. Furthermore, we found that often the more important the variable in determining an outcome, the more difficult it is to quantify, thus limiting the usefulness of numerical analysis **by itself.**

Table 1 is a synopsis of the characteristics and findings of eight of the studies we have analyzed using primary data. Each data set has its own unique properties and relative strengths and weaknesses. In all instances there was some evidence of favorable change in the service area. However, in three of the eight cases, this change was most likely due to factors not related to the intervention. In the fourth, our confidence in attributing the change to the intervention was weak. Volumes 2 through 6 report these case studies.

The judgment, experience, and "feel" of project personnel and others was absolutely essential in performing, interpreting, and analyzing the numerical data. In no case was the research team able to fully understand its analysis until it had been reviewed by people in the field. Even though at least one person from the research team had visited each intervention site prior to analysis (and in several cases spent considerable time learning about the intervention and its environment), only field people could truly understand why trends in the data appeared. Successive cycles of analysis, review, and further analysis were necessary to reduce the array of competing explanations to the point where it was possible to incorporate the project staff's judgment and arrive at a satisfactory conclusion.

C. The Consequent Approach to Analysis

The aforementioned methodological findings caused the research team to evolve a philosophy and approach to evaluating that resolves many of the issues in a useful manner. In the final year of the project, portions of this methodological approach were tested by application to several of the interventions studied in depth. The approach can be summarized as follows:

TABLE 1

SYNOPSIS OF CHARACTERISTICS OF SELECTED PROGRAMS USING PRIMARY DATA

	Name & Location	Approximate Population Observed & Period of Data	Observed Outcome	Possible Attributions or Explanations for Outcome	Most Likely Attribution(s)	Confidence of Attribution	Type of Control Utilized
1	Candelaria, Colombia	1200 families (1968-1974)	• Malnourishment dropped from 29% to 21% (25% reduction) • Diarrhea prevalence decreased	• Program effects • Improvement in environment • Reduction in disposable income • Aging population • Target population change • Selection biases	Promotora Program effects	Strong	Reflexive, statistical
2	Candelaria, Colombia (Revisited)	500 families (1976)	• Reduced diarrheal disease • 30% reduction in pregnancy rate	• Program effects • Target population change • Change in environment	Synergism between Promotora Program and other programs	Strong	Participants & nonparticipant groups (statistical)
3	Kottar Social Services Society Tamil Nadu, India	4000 families in 21 villages (1975-1977)	• Malnourishment dropped from 50.2% to 42.5% (15.3% reduction)	• Program effects • Environmental changes (drought) • Aging population • Selection biases	Program effects and/or improved environment	Strong	Reflexive, statistical
4	Esperanca Santarem, Brazil	460 children in 4 villages (1977-1979)	• Malnourishment dropped from 48% to 33% in program village (31% reduction)	• Program effects • Economic conditions improved • Environmental conditions • Aging of population	Program effects combined with improved economic conditions	Moderate	Treated & untreated communities, statistical
5	Primops Cali, Colombia	11,700 children (1976-1977)	• Malnourishment dropped from 31% to 27% (13% reduction)	• Program effects • Change in target population • Improvement in water & sewer • Population aging	Program effects combined with improvements in environment and target population	Moderate	Reflexive, statistical
6	Rice Fortification Project Chang Mai, Thailand	1200 children in 29 villages (1971-1975)	• Slight reduction in malnourishment	• Economic improvement • Environmental changes • Program effects	Environmental changes	Moderate	Variously treated communities, reflexive, statistical
7	Community Development Project Tangse, Indonesia	500 children in 11 villages (1977-1979)	• Slight reduction in malnourishment	• Aging population • Program effects • Environmental conditions changed • Selection biases • Target population change	Aging population combined with program effects	Weak	Reflexive, statistical
8	Community Development Project Pespire, Honduras	500 children in 11 villages (1977-1978)	• Malnourishment dropped from 28.9% to 23.1% (20% reduction)	• Target population change • Selection bias • Drought condition changes • Program effects	Selection biases combined with possible changes in target population and environmental conditions	Moderate	Reflexive, statistical

1. Recognize at the onset that except in rare instances, it will not be possible to build a completely solid analytical conclusion regarding the causes of a change in the target population, whatever the experimental design.

2. Given this indeterminacy, base an evaluation and feedback scheme upon a realistic appraisal of: (a) the resources and financial constraints on the project, and (b) the best way of building evidence for attributing the outcome(s) to a cause. This approach certainly involves experimental design issues, and data gathering and processing consideration, but also includes schemes for incorporating less quantitative and more judgmental factors as well.

3. Plan at the onset for successive cycles of analysis and project staff review. Modification of the data collection scheme and/or types of analysis should be made during these periodic reviews.

III. SUBSTANTIVE FINDINGS

A. Findings with Respect to Process

The substantive findings of the research team reflect many of the same issues that affected the methodological study. The same issues that confound analysis of data also confound efforts to run a good nutrition intervention. Several general conclusions relating to all of the interventions studied act as a starting point for discussing findings:

1. The setting in which a community-based intervention is carried out will most likely change in major ways during the course of intervention.

2. While the state of the art of community-level problem diagnosis has improved recently, it offers at best only a first approximation to an understanding of the specific causes of malnourishment at the local level.

As a consequence of these two observations, the following conclusion can be drawn.

> It is virtually impossible to derive, in advance, an intervention strategy capable of producing optimal results. At best, one can arrive at a good initial plan. A strategy

invoking mechanisms for modifying the intervention as time progresses is needed to enable successive shifts toward the optimal intervention in response to the continually changing program setting.

If one accepts the notion that a good intervention is one that builds in mechanisms for self-evaluation and improvement, one must then articulate which components of an intervention contribute to its ability to change in a positive way. These are:

- If the designers of an intervention articulate a theory of the causes and effects of malnutrition in their communities, they are more likely to identify viable strategies for combatting that problem. The theory provides a framework for selecting program components and assessing their impact.

- A nutrition program capable of monitoring and evaluating itself is more likely to reduce malnourishment than one that does not have this capability. If the environment changes and first solutions are approximations, data gathered on-scene serve as the best sources to guide informed program modification.

- While sophisticated analytic schemes are sometimes helpful, simplicity at the onset is always a virtue. The use of simple two-way contingency tables wherever possible facilitates participation of project personnel who may not be versed in the technologies of numeric assessment. Graphic representations of the relationship between malnourishment and age in a population, which we call characteristic curves, should be generated to portray nutritional status in a manner which incorporates the changes in nutritional status due to aging of the child. These characteristic curves were used in the early stages of each of our analyses and were found to be easily understood by field-level personnel. In addition, they were found to be readily constructed under field conditions, thereby facilitating speedy analysis.

- Nutrition interventions that emphasize local involvement, both in design and implementation, are more likely to succeed than ones that do not. Local people are most capable of understanding their environment and can sense what will work in that environment. Furthermore, support of the local infrastructure is often critical for success, and support usually follows involvement.

- A nutrition intervention director who attempts to promote linkages with nonlocal systems raises the likelihood of reducing malnourishment at the community level. The additional resources derived through such linkages--knowledge, money, and materials—can greatly facilitate the implementation of locally derived plans.

In summary, successful community-level intervenors embrace a process that uses an incremental approach to problem-solving, whereby the evaluation of outcome information and process data guides the choice of successive increments. Furthermore, those same successful intervenors use local resources and involve local people as much as possible while retaining viable ties to the outside world to keep valuable resources flowing into the community.

B. Findings with Respect to Content

Thus far, the findings of the "analysis" project concerning components of interventions have dealt with process--not content. Most casual observers might be surprised by this. Analysis or evaluation is usually taken to mean "Does a latrine program improve nutritional status?", or "Does take-home food supplementation work better than food supplements given at a central place?" This selected emphasis on process is by design. Problems in different locations are different. Personalities and skills for different interventions are different. There can be no absolute statement about what works, because of the importance of the local context.

Since there are such complex local and regional factors that determine the efficacy of a particular program or program component, we favor an approach that draws upon the current literature and professional expertise to formulate possible intervention designs--but only as a first approximation. Care in drawing upon previous work should be taken to ensure: (1) the comparability of the settings, and (2) an assessment of the amount of evidence in support of a causal link between the observed outcome and the intervention. Often if the aforementioned competing-explanation methodology is rigorously applied, ambiguities become more apparent than as reported in the reference documents.

Still, some things can be said about specific issues. Chapter VI discusses uses of food supplementation, health (both preventative and curative), environmental conditions, and family planning components.

IV. A RAPID INFORMATION FEEDBACK SYSTEM (RIFS)

There is a surprising congruence between the methodology proposed for analysis and the process proposed to enhance the probability of creating a successful community-level intervention. Both call for ongoing longitudinal data collection. Both call for periodic analysis of data by both the analysis team and the field staff. Both call for modifications as more is learned about data needs and competing explanations of outcomes. The distinction between the two (it is not a difference, merely an extension) is that the self-monitoring and evaluation process for program planning incorporates the results of the analysis into action in the field, not only in published reports.

An interesting consequence of the decision to utilize data analysis to guide programmatic decisions is that the data are of higher quality and the analysis is subject to greater scrutiny. A "live" data set, in the sense that its analysis will contribute to program decisions that affect the data gatherers, will locate sources of errors and create a reason for the elimination of those errors.

We have come to call such a "live" data gathering and analysis system, placed in the context of program planning and management, a Rapid Information Feedback System (RIFS).

In summary, the following is a list of steps to be taken to implement a Rapid Information Feedback System:

1. Develop a logical construct of the "nutrition system" in a community and design a data collection system that captures key variables.

2. Prior to initiating the intervention, develop a system (set of programs) to clean and validate data--preferably one that can be operated as close as possible to the source of the data in both time and place.

3. Analyze the data:

   a. identify competing explanations for observed outcomes from preliminary analysis;
   b. use simple and/or more complex analysis to eliminate implausible competing explanations;
   c. review analysis results with field staff.

4. Modify both the data system and the intervention in response to insight gained from analysis.

5. Iterate steps 3 and 4, using an ongoing data collection process, recognizing continuously that while local intervenors may not have a worldly perspective, they do have the rich substantive knowledge of their community that must be recognized and utilized.

CONTENTS

CHAPTER V

TABLES

FIGURES

MAPS

CHAPTER I
INTRODUCTION

PROJECT OBJECTIVES

Assisting developing countries through nutrition intervention programs is a very serious undertaking: people's lives are at stake. Perceiving that to improve the quality of assistance provided requires a continuing reassessment of nutrition interventions, the Office of Nutrition of US/AID initiated an effort to learn more about which aspects of interventions work well, and which do not. Community Systems Foundation accepted this charge to examine the experience of others and thus offer findings useful for improving present and future programs of nutrition intervention in developing countries.

In accord with these goals, a project was defined with the following objectives:

1. to identify characteristics of community-based nutrition interventions which are associated with success;

2. to report these findings in a way which will assist others in raising the likelihood of success; and

3. to recommend ways in which national surveillance systems can draw upon locally generated data.

It soon became apparent that to achieve these objectives, we would have to add a fourth:

4. to design and implement a system capable of providing rapid feedback of analysis to ongoing interventions.

PROJECT SCOPE

The Definition of Terms

The definitions of the key terms in this report play an important role in establishing the limits and nature of its content.

1. Analysis

Although the conclusions of this report are based, in part, on a non-quantitative review of a great many nutrition programs, the primary effort was directed toward the application of a variety of quantitative techniques to numeric data related to nutrition and its indicators. Because "analysis" was taken to include such a strong quantitative component, programs accepted for review had to be engaged in the recording of numeric data.

2. Community-Level

Nutrition interventions rarely emanate from or subsist solely upon community resources alone; each combines local and non-local resources to varying degrees. Yet, rather than simply being local manifestations of nationally defined programs, nutrition projects accepted for this analysis had to be operating at the community level; that is, they had to have distinct locally defined and implemented components. Policymaking authority, with regard to the choice of activity and method of implementation, had to reside in part with the individual community.

3. Nutrition Interventions

Many nutrition projects are designed to improve food availability through supplementation, fortification, and increased production. Others strive to improve health status through hygiene, nutrition and family planning, or through improvements in potable water, waste disposal, and health care delivery systems. Still, defining nutrition interventions in these terms alone ignores other and possibly more profound determinants of malnutrition, such as inequitable income distribution and land tenure. Yet for this research project to be useful, it had to restrict its domain to manageable proportions while at the same time not lose sight of the larger context. To avoid definition in terms of specific activities, this analysis describes nutrition interventions as projects that: (1) explicitly state

improvement in nutrition of preschool children as one of their goals, and (2) focus efforts on resolving one or more problems directly associated with malnourishment.

4. Success

The most obvious means of defining a successful nutrition intervention is in terms of the nutritional status of the participants; if nutritional status improves as a result of the intervention, it is a success. In the analysis, improvement in nutritional status over the life of an intervention was taken as the primary indicator of success. Yet simply measuring change in nutritional status during an intervention does not indicate whether a permanent change has occurred. Unfortunately, short-term improvements in nutritional status frequently come at the expense of creating dependencies on external resources. This dependence may later adversely affect the community. Even more clearly, temporary changes do little to ease problems of chronic malnourishment. Although long-term improvements in nutritional status may not become immediately evident, only permanent change can hope to alleviate worldwide malnutrition, and we believe success must be judged in this regard. Therefore, this analysis focuses upon nutrition interventions that: (1) have the potential for continuation after the pilot stage, and (2) do not create heavy dependencies that make communities unduly vulnerable to shifts in political forces.

In summary, the scope of this report is limited to interventions with a distinct locally defined component, an explicit goal to improve the nutrition of preschool children, potential for long-term continuation beyond the pilot stage and a numeric data base amenable to quantitative as well as qualitative review.

The Need for In-Depth Studies

Quantitative probes of the success of an intervention can be pursued at varying levels of intensity. With few exceptions, the analyses described in the literature are cursory—summaries of the changes observed in selected

indicators of nutritional status. Due to the complexity of the relationships among the many factors which contribute to the nutritional status of an individual or community, such cursory analyses can be dangerously misleading. They tend to ignore nonprogram events such as drought, famine, inflation, rain, harvests of plenty, increased employment opportunity and many others. These must be adequately understood and accounted for in order to enable proper interpretation of observed changes in nutritional status. Consequently, this research emphasizes in-depth analysis of a relatively small number (8) of interventions, rather than a cursory review of a larger set.

In this regard, this research is highly unusual, if not unique. Surprisingly few interventions have been subjected to the objective analysis of change in nutritional status in designated target populations. Of nutrition programs surveyed in a recent study by the Harvard Institute for International Development, less than one-quarter gathered and analyzed nutritional data at **any** level (Austin, 1978). A survey of 180 health programs in the developing world conducted by the American Public Health Association reports similar results; few health programs that had a nutrition component analyzed data taken on the target population (Karlin, 1977).

Previous research that does examine the effectiveness of nutrition interventions can be classified in two general categories: (1) analysis of cross-sectional data drawn from several interventions, exemplified by the Checchi evaluation of supplemental feeding programs (Checchi et al., 1974); or (2) analysis of time series (longitudinal) data from single interventions, best illustrated by the Drake, Fajardo study of the Promotora program in Candelaria (Drake and Fajardo, 1976). Strikingly absent from prior research has been a comparative analysis of several nutrition programs collecting longitudinal data. This project represents one of the first research efforts of this nature, and reflects a commitment on the part of the sponsoring agency to develop a firmer understanding of the characteristics of externally funded nutrition interventions that contribute to a success.

PROJECT EVOLUTION

The balance of this report is a series of findings which summarize the knowledge derived by its authors during four years of research. Coming as it does at the end of the project, the report is a "snapshot" of both an approach to analysis and a philosophy of intervention. To place the findings with regard to both the methodology and the philosophy in perspective, it is useful to look briefly at the evolutionary process which led to this "snapshot."

But first, it is important to remark on a singular peculiarity of this evolutionary process. At the outset, two distinct questions were asked. The dominant question, given the objectives, was what makes for a successful nutrition intervention. To answer that question by reviewing ongoing interventions, we first had to ask how does one determine whether an existing program is successful. A single answer emerged to both questions: the process by which program results are accurately determined is precisely the same process which enhances program success.

The initial response to the second question on methodology was the traditional response to the individual case-study approach. Data sets were to be gathered that contained similar, if not identical, yardsticks with which to measure success. The elements of the interventions generating the data sets along with other data describing the settings of those interventions would be quantified and, using multivariate statistical techniques, related to the success measures. This approach is exemplified by the Morss et al. (1976) study of small farmer development. Their analysis attempted to explain variance in success measures by "independent" variables that described the interventions. Both the measures and "independent" variables were articulated in a model of rural development that served as the guide to data collection. For a variety of reasons, the methodology giving rise to this report diverges from this intuitive approach.

1. Inadequacy of Conventional Statistical Analysis

The Morss et al. study is not only an example of a traditional analysis across several interventions, it is illustrative of many of its problems. Realizing the need to account for non-quantifiable factors that affect interventions, Morss et al. constructed complex indices from questions requiring the subjective response of their subjects. These served as their "independent" variable. This process is elaborate—almost mystical—and tends to mask the weakness of the double use of subjectivity (first by the respondent in providing raw data and second by the analyst in manipulating that raw data to derive indices in the appropriate form). The conclusions derived from the multivariate analysis using these indices are presented in the text as "high-science" while the magical manipulations done prior to the statistical work are buried in appendices.

We too recognized the need for including subjective variables in our analysis. One cue, in fact, comes from Morss et al.

> At this point we offer a cautionary note that should be borne in mind throughout the reading of this report. While we believe our quantitative approach does reveal new insights, it does not offer any major formula that will guarantee success. Rather, quantitative analysis is seen as an imperfect but useful tool to complement the common sense, patience and the thorough knowledge of local conditions that are essential for persons designing and implementing rural development projects for small farmers. (Morss et al., 1976 p.19)

The footnote in the above citation went on to urge the reader to read the detailed accounts of each case study to draw other important lessons not obvious in the statistics.

Our response to this need to incorporate the "non-quantitative" in the analysis was to learn as much as we could about each intervention and to factor that knowledge into the interpretation of each separate and distinct analysis. In so far as parallels existed in the separate analyses, generalizations could be drawn.

2. Diversity in the Nature of the Data Sets

In fact, such generalizations were hard to come by. One reason for this is the very diversity of the data sets included in this analysis. (Whereas traditional studies examining multiple cases generate their own data in a form consistent across all cases, this study relied on data collected by each intervention in the format most appropriate for use by the intervention.) The only elements common to all our data sets were the ages and weights of preschool children. The data were collected at different intervals (one month in KOTTAR to six months in PRIMOPS), spanning different time periods (15 months in Honduras to eight years in Candelaria), for different numbers of children (589 in Indonesia to 12,770 in PRIMOPS). The composition of the data sets was equally diverse (complete medical examination including hand x-ray in Thailand to a fairly exhaustive socioeconomic profile for the family in Esperanca).

Similarly, the eight interventions showed considerable diversity with regard to program, level of intensity and, even, purpose. (KOTTAR relied on monthly home visits, food aid, and viewed their health program as part of a larger development effort, while Esperanca established health posts, provided education programs, and trained University students in rural medicine.)

3. Volatility of the Total Program Environment

The diversity in the composition of the data sets and the make-up of the programs was not the only source of variation making the traditional research difficult. The setting for each intervention was unique and, more importantly, each setting exhibited a volatility of its own.

With regard to any social intervention, volatility of the program setting is a potential hazard for scientific research. This is especially true in the international nutrition field. Although the intuitive solution to poor nutrition--increased consumption of proper foods--is still the central element of nutrition intervention, the importance of disease in negating the effects of improved consumption is now widely acknowledged by nutrition planners. Still, interventions directed at only improved consumption and/or

reduced disease are, in fact, operating at only one level of a very complex system. The underlying economic, political, social, and physical components of the total environment do more to determine overall consumption and disease rates than can any intervention. If these components remain stable, analysis can proceed according to traditional precepts—but they never do.

In response to these perceived problems with conventional research strategies, we turned to an approach that acknowledged their existence. To the degree possible, we performed parallel, in-depth case studies of the interventions selected for analysis. These case studies embraced two highly unusual tactics.

First, to account for the unique character and environmental setting of each program, we encouraged each intervention staff to participate in the interpretation of the analysis of their own data. Although we recognized that having project personnel serve an evaluative function allows for the introduction of an analytic bias, we felt that the intimate knowledge of the total environment held by people on-site was critical to our understanding of the analysis.

Second, to facilitate the identification of the confounding factors in an analysis of social systems, we explicitly sought all competing explanations for observed outcomes and tested their plausibility using the subjective opinion of the project staff as well as the results of statistical analysis.

To implement these tactics, it was necessary to develop the capability to analyze data rapidly and to return the results to the field for review. We came to call this capability Rapid Information Feedback System (RIFS). The RIFS concept encourages field people to identify competitive explanations of trends in the data, assess their plausibility, and otherwise participate in the anlysis. (The RIFS concept is developed more fully in Volume VII.)

The RIFS concept evolved in response to the second question posed earlier—how does one determine whether an existing program is successful? Application of the RIFS concept taught us an invaluable lesson with regard to the primary question—what makes for success? The programs that were

both amenable to and capable of learning from their own review were also those which appeared to be successful.

ORGANIZATION OF THE REPORT

We now turn toward the evidence for this last conclusion. Because the conclusion emerged in response to the search for an effective analytic strategy, much of this evidence is a discussion of the methodological issues encountered during that search.

We will begin in Chapter II with a brief summary of each intervention and the associated data analysis. More detailed accounts are included in Volumes II through VI of this report. Because of the centrality of the concept of nutritional status in the analysis, Chapter III discusses various findings with respect to the measurement of malnourishment. This chapter reveals the strengths and weaknesses of using anthropometrics to measure success.

One factor that contributed heavily to the growth of the RIFS concept was the poor quality of the data analyzed. Chapter IV describes our efforts to clean the data and, in accordance with our second objective (to report our results in a way to assist others), sets out some guidelines to help others improve the quality of their data.

Chapter V is a discussion of commonly encountered problems in analyzing nutrition data and comments on our "solutions" to those problems. Finally, with our discussions of the theoretical, data, and methodological problems confronting nutrition analysts, we turn to our prescription for successful nutrition intervention. This makes up Chapter VI.

CHAPTER II
THE DATA

The data examined in the project consist of "hard" data (anthropometric measurements supplemented in some cases with other socioeconomic and/or health data) for ten distinct interventions, and "soft" data (our observations on the process employed) for those same interventions. The two types of data are inextricably bound together; each makes sense only when considered in the context of the other.

The "hard" data offered us the opportunity to test for and quantify changes in outcome variables (changes in nutritional status and/or health indicators) and to relate those changes to the other variables in the data set. The "soft" data helped us interpret the analysis of the anthropometrics in terms of the process of the interventions. To understand those processes, we constructed a framework for describing interventions based on a concept of project life cycle. Attempts to generalize on the development of an intervention have led to identification of typical stages in most interventions. Because our written descriptions of the interventions will be in terms of this life cycle framework, it will be described briefly here.

We begin by identifying four stages in the life cycle of an intervention:

1. Entrepreneurship
2. Community Embedding
3. Durability At Existing Scale
4. Scaling-Up

These stages are not necessarily separate and distinct; there is often considerable overlap and, in some cases, a complete reiteration through some or all of the stages.

In general terms, the "entrepreneurial" stage begins when some outside change agent builds an alliance with a community, or a community pulls an

outside agent into itself. "Embedding" is the planting of the idea (proposed solution) behind the intervention in the community and the nurturing of the idea until it gains acceptance. Often, the idea is modified during this stage to enable embedding to occur. A negotiation between the community and the change agent to match resources to perceived needs is not uncommon. The "durability" stage is the one in which the intervention is actually effectuated in a more-or-less stable way. Finally, "scaling-up" is the expansion of the intervention either geographically or in the scope of services. Few interventions remain confined to the originally designated recipient population or stay with an original package of services.

An important concept underlying this life cycle model is the notion that the change agent acts as a bridge between two worlds—the world of the community and the outside world of international finance and knowledge. Both worlds are constantly changing. The entrepreneurial change agent is forever trying to bring the resources of the outside world to bear upon the changing community in a useful way.

The "bridging" notion suggests two potential forces that move an intervention through its life cycle: changes in the change agent, and changes in the resources he/she can command. The first force, the changing personality of the intervenor, is critical to understanding the life of an intervention; the characteristics of the intervention are so often a reflection of the character of the primary change agent. Actual shifts in personnel and/or changes in the aspirations and interests of continuing personnel account for many of the transitions between life cycle stages.

Similarly, the second force, changing resources, accounts for many of the transitions between stages. In particular, the exhaustion of initial resources over time often prompts scaling-up in the form of an enlargement of the recipient group and/or the introduction of new services so that "new money" can be found to keep the intervention going at all. We will argue throughout this report that data analysis should be a third active force in causing changes in interventions. One characteristic of the interventions considered in this project is that some data were collected

for each. Our first major observation is that in almost no case was the data used to help shape the intervention.

A pictorial representation of the life-cycle process is presented in Figure 2-1. The stages of the intervention are illustrated. The flow of time is demonstrated by the large arrows, and the bridging notion is conveyed by placing the intervenor between the two worlds.

The balance of this chapter is dedicated to a description of the interventions and data sets studied during the course of the project. Each description will attempt to place the history of the intervention in the context of the life cycle paradigm, giving explicit attention to the personality, resource, and data analysis forces contributing to change. Finally, a summary of the results of the analysis will be provided. There are seven interventions with longitudinal data plus an eighth data set arising from a revisit to one of the intervention sites two years after the phasing out of services. In addition, we have two cross-sectional data sets which were used to help test some of our methodological procedures. Map 1 indicates the location of all of the interventions on a world map.

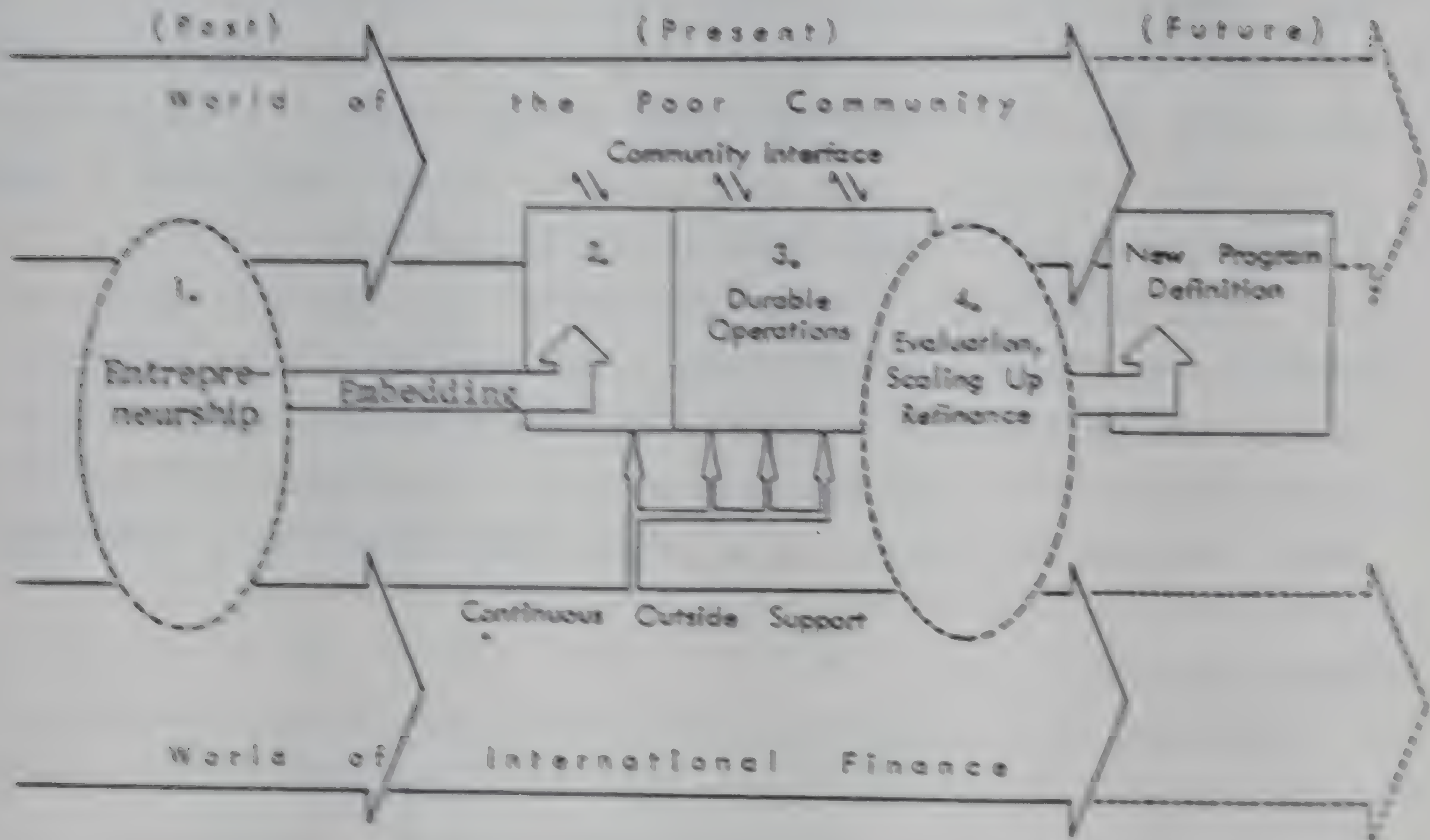
ESPERANCA

Esperanca is a private, non-profit organization whose primary goal has been the provision of basic health care to a population living in northern Brazil's central Amazon River basin (Map 2). Because of marked discontinuities in the tenure of the primary change agents who have contributed to the Esperanca effort, the intervention has passed twice through the life cycle stages and is currently sustaining a high level of activity, including an expansion of services to many more rural villages in the region.

The original program was initiated by the late Father Luke Tupper who, as the only priest-doctor in the Franciscan Order, traveled in the Amazon by whatever means he could devise, providing medical care to the isolated river communities. His recognition of the need for outside resources, particularly to establish a massive immunization program, prompted the formation of Esperanca, Inc., in Phoenix, Arizona, in 1970. In 1972, Father

FIGURE 2-1

THE INTERVENTION PROCESS



MAP 1 - PROGRAM STUDY AREAS

STUDY AREAS: ✪ Longitudinal Data

★ Cross Sectional Data

• Other Sites

Luke opened a medical clinic in Santarem on land donated by the Franciscan Order, and Esperanca bought a boat which was converted to a medical hospital and brought to the Amazon in May 1974. Father Luke returned to the United States at the end of 1975 and the project was essentially dormant until January 1977.

At that time, under a new director, Dr. Harry Owens, the program experienced a rebirth. The medical clinic reopened under the direction of Dr. Fred Hartman. This program rebirth included the addition of a more "preventive" program component for Esperanca. This component stressed early treatment of diarrhea, oral rehydration, reduction of parasites, regular weighing of children and/or home visits of children under six, and the holding of general education programs. A nutrition rehabilitation center was opened in Santarem to provide intensive care to third-degree malnourished children who "lived" at the center until they recovered. An ambulatory program for other malnourished children and their families was also established. Finally, an outreach program in two communities, Mojui dos Campos (pop. 3,000) and Alter do Chao (pop. 800), was started. These two communities were chosen because of community-initiated requests for Esperanca medical services. The outreach programs were similar to the ambulatory program in Santarem—consultations with a visiting doctor or the trained "barefoot" doctor manning a health post, regular child weighings, and an education program.

In June of 1978, Dr. Hartman left Brazil. Dr. Owens was also no longer on-site. As a result, the outreach programs--the Esperanca component of most interest to this project—went into a "durability" stage characterized by stagnation. Particularly in Mojui dos Campos, the larger and more poorly organized village, the activities of the health post were sporadic and infrequent. In the smaller village, Alter do Chao, the program continued to operate as anticipated.

One year later, Dr. Owens returned to Brazil with a public health specialist, Steve Alexander, to pump renewed life into the existing program. The resources for this activity come from a grant from PACT to "scale-up" the outreach program to cover eighteen villages over the next

two years. As of the time this report is being written, the second year of activity under PACT funding is beginning. Additional health posts have already been opened and the training of more health workers is progressing well.

Esperanca - Data

As part of its initial outreach program, the Esperanca staff administered an extensive epidemiological survey of all children in Mojui dos Campos and Alter do Chao in 1977 and 1978. The survey was taken for three reasons: (1) to obtain baseline data for diagnosing community problems; (2) to obtain baseline data for measuring program impact; and (3) to identify and describe communities to serve as a "laboratory" for students at the University of Para.

The survey was administered in four villages—the two program villages and two other "control villages." Initially, interviews were done in Mojui and the two control villages. The request for Esperanca services from Alter do Chao some six months later led to its inclusion in the program. This inclusion of Alter do Chao as both a program and survey site resulted in a neat experimental design pairing two inland villages, Mojui and Castanhal, and two coastal villages, Alter do Chao and Aramanai. The intention of the Espananca staff was to resurvey the same villages two more times at intervals spanning five years.

The first resurvey, conducted in May of 1979, was jointly administered by Esperanca staff and staff from Community Systems Foundation (CSF). Eighteen months had elapsed between surveys for three of the four villages, while only twelve months had passed for Alter do Chao. The resurvey questionnaire contained more detail than the original; however, care was taken to preserve comparability between the two. Both included socioeconomic, family history, medical, and anthropometric data.

The data were keypunched and entered directly into the data base management system, MICRO. The MICRO system facilitates the merger of survey and resurvey data into a single data base for comparative analysis, and will enable Esperanca to add to the data base easily in the future.

Errors were removed from the data by using the facilities of the data base management system. Variables with "illegal" codes could not be entered into the system until they were corrected, and consistency checks using the MICRO command set were done to locate inconsistent entries in related questions. Finally, subsets of the data set were constructed using MICRO. These were passed to the statistical system, MIDAS, for analysis.

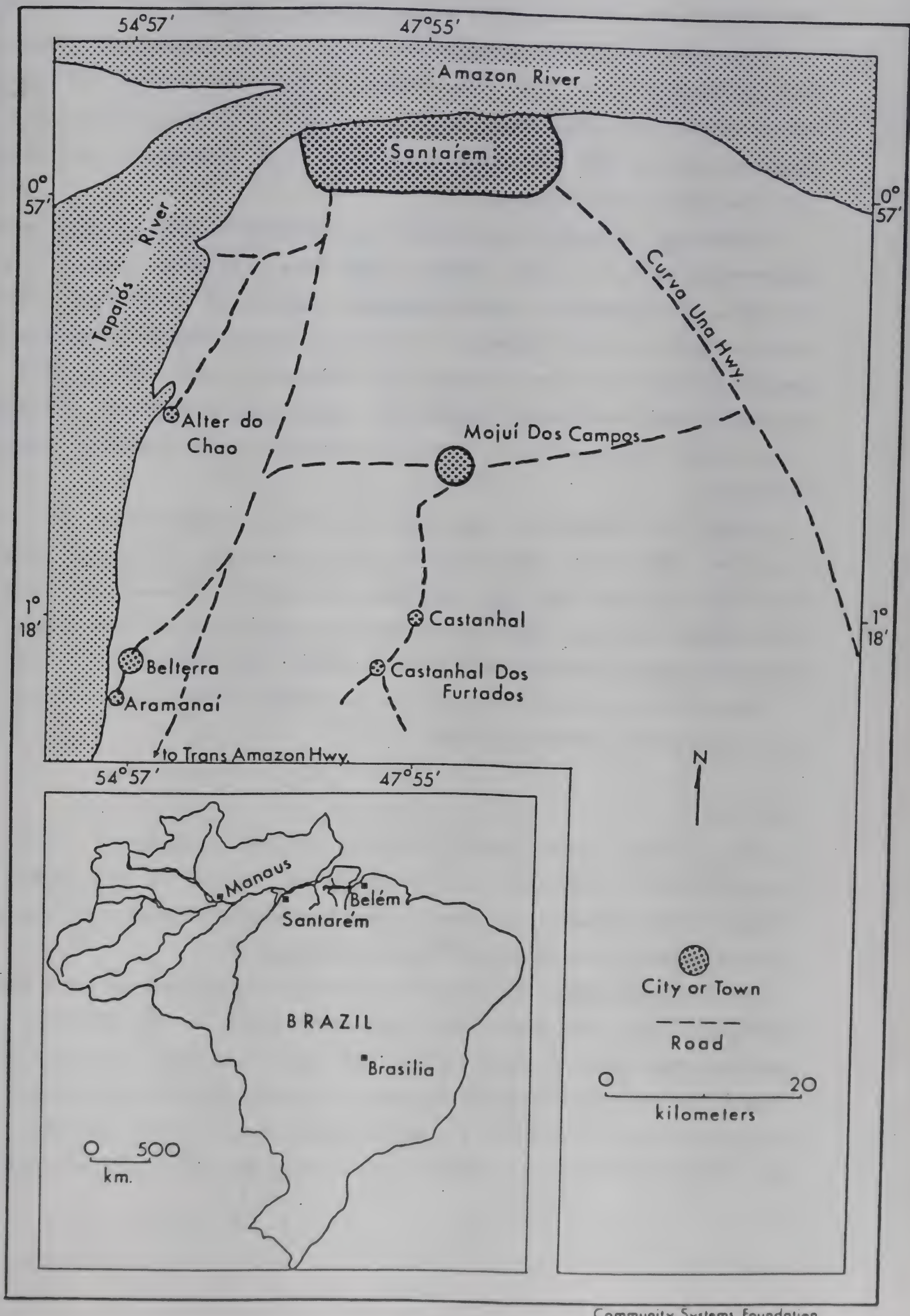
Esperanca - Analysis

The analysis of the Esperanca data sets illustrates the difficulties involved in interpreting the results of a social experiment even when that experiment was implemented using a classical "laboratory" design. In this instance, the classical matched-pairs design was only approximated; the villages were neither randomly assigned to groups nor matched by any careful analytic procedure. Furthermore, the elapsed time between observations was long enough to allow considerable variations in nonexperimental factors in the target populations. There was in- and out-migration in the villages, children became over-age, and the socioeconomic environment changed unequally in the four villages. Finally, the intensity of program application was not sufficiently well documented to enable clear attribution of observed outcomes to program components.

The analysis concentrated on identifying change in selected indicators of health. These indicators were selected because they reflected explicit goals of the Esperanca outreach program. Overall nutritional status was measured, using weight as a percentage of the Gomez weight-for-age standard. Anemia was measured using hematocrit readings. The prevalence of parasites was measured by identifying the presence or absence of particular parasites during a fecal examination.

The use of weight as an indicator of growth led to mixed results. Whether all children were compared or just those who participated in both surveys, the pair of inland villages--Mojui and Castanhal--showed counterintuitive outcomes. The program village showed a decline in overall nutritional status, while the control village improved. Where the program was more dynamic, on the coast, the result was anticipated. There was

MAP 2 - ESPERANCA PROGRAM - SANTARÉM, BRAZIL



improvement in Alter do Chao, while Aramanai remained much the same.

With regard to anemia, there were relatively few non-normal hematocrit readings--only about 10%. The children who were deficient in the first survey recovered, while a new set of children of the same age who had been normal for the first study showed signs of anemia in the resurvey. This was true in all villages.

Comparison of parasite prevalence is difficult because of the seasonal factor—the surveys in three villages were done 18 months apart. In Alter do Chao, the village of intense program application, the surveys were 12 months apart, and the number of children with ascaris infestation (a serious problem in that village) had declined from 47% to 22%. In Aramanai, over an 18 month period, the percentage of children with ascaris infestation was constant. Similar consistency was found in Mojui and Castanhal.

Overall, the village of Alter do Chao showed improvement on two of the three indicators. Because this was the village of highest program activity and was the only one showing such improvement, there is substantial evidence that, at a high enough level of intensity, the Esperanca program has reduced malnourishment and disease.

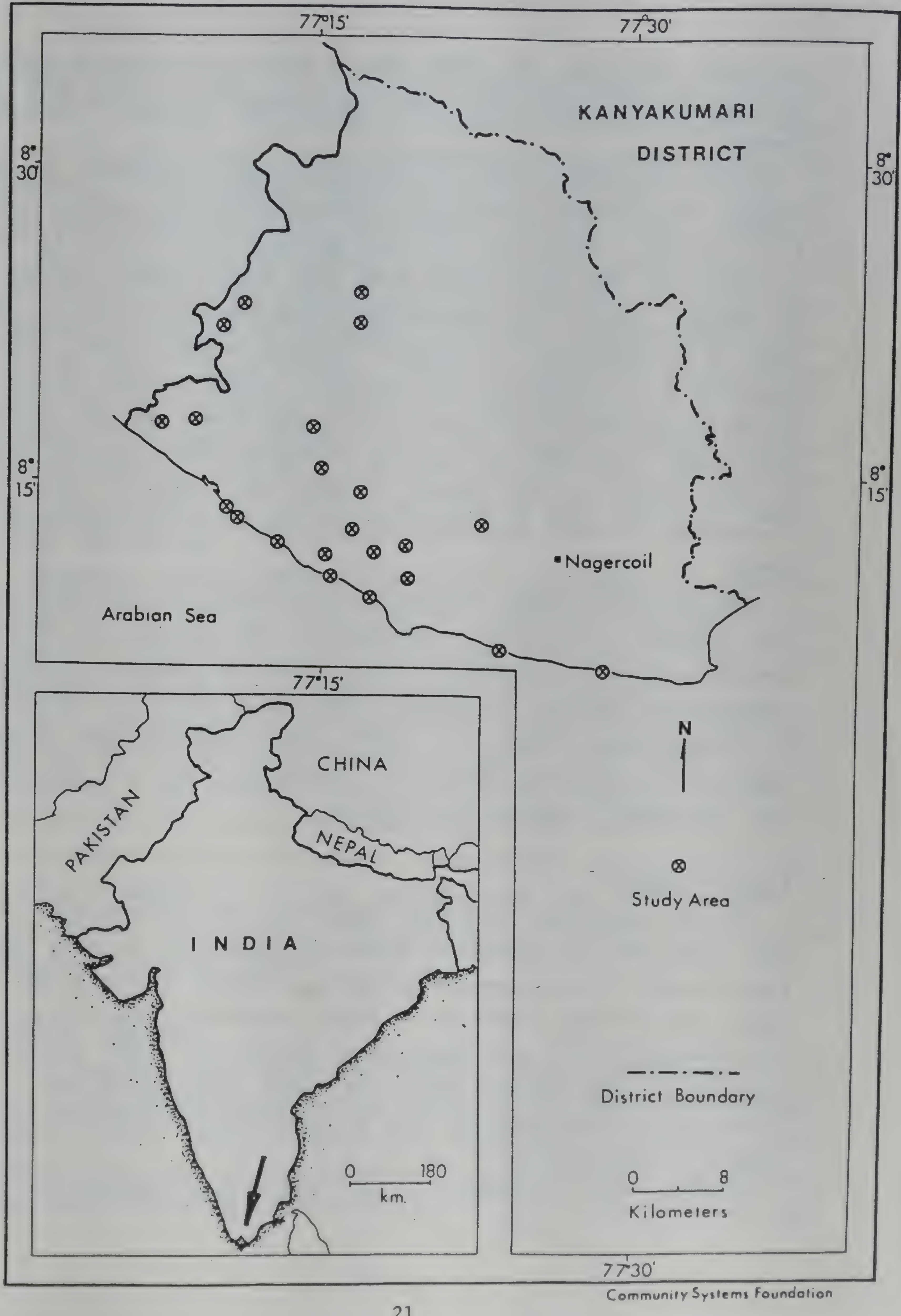
Volume II of this series is the complete study concerning the intervention and outcomes to date.

KOTTAR

The Kottar Social Service Society (KSSS), founded in 1962, is a volunteer agency affiliated with the Catholic Church in the diocese of Kottar. This diocese is located in the Kanyakumari District of the State of Tamil Nadu in the southern tip of India (Map 3).

The entrepreneurial drive behind KOTTAR is the presence of a Belgian priest and sister who make their permanent home in the district. As a result of this primary identification with the study area, the pace of the project has been slower and community involvement has been greater than in projects where the primary change agents devote only a portion of their lives to the intervention. Much of the time since 1962 has been spent in

MAP 3 KSSS PROGRAM - KANYAKUMARI, INDIA



community embedding. The KSSS tries to work out strategies to promote development with villages which show some interest in working with the Society.

In 1972, the KSSS created its Community Health Development Program (CHDP), which includes a take-home supplementary feeding program, health and nutrition education, and some primary health care, including immunizations. The CHDP covers more than 38,000 preschool children in 124 villages. Many of the participating villages have participated in other types of interventions. A potters' cooperative was formed, and an outgrowth of its financial success has been a potters' housing cooperative. Fishing villages have experimented with larger boats and net weaving cooperatives. The KSSS has attempted to organize people to plant cactus in areas subject to soil erosion, to build tube wells to level out fluctuations in water availability, and to reconstruct roads destroyed by floods.

There has been concern on the part of the founders of the project that their presence is too critical a component of the intervention; that is, durability will not be achieved until local people are trained to provide the organizational skills and the bridges to the outside world now provided by the Belgian Clergy. (Many of the KSSS activities utilize outside resources found by the Father and Sister in the developed world.) At this time, they are working to achieve this durability.

Kottar - Data

The data describing the Kottar intervention come from a survey done with funding from the Rockefeller Foundation and CSF by Dr. John Osgood Field, formerly of the International Nutrition Planning Program at MIT and now at the Nutrition Institute of Tufts University. The survey was administered in 1977 and consisted of health and weight histories of approximately 4,000 children (taken from weight cards kept as part of the CHDP) and a cross-sectional survey of socioeconomic information on the families of these children.

The data were keypunched at MIT and a tape containing the card

images was provided by Dr. Field to facilitate analysis at Community Systems Foundation. The process of preparing the data for analysis proved to be far more complex than anticipated due to keypunching and coding problems. It was therefore especially helpful to be able to utilize the data base management system for facilitating the next stages of analysis.

First, using a unique FORTRAN computer program written especially to perform these tasks, the longitudinal data were extracted from the surveys, cleaned, and expanded by the addition of several variables computed from the raw data. These variables included the percent of standard scores using the local standard, the Harvard standard, and the NCHS-CDC standard. The cleaning process reduced the number of records containing at least one error from 25 percent to 2 percent. The final data file contained weight and health time series for 4,075 children—a total of 79,687 observations.

Similarly, special FORTRAN programs were compiled to process the cross-sectional portion of the data set. Whereas 40 percent of the interviews contained some error at the start, again only 2 percent of the interviews could not meet the test for errors after cleaning.

Both the longitudinal and cross-sectional data files were incorporated into the MICRO data base management system. To reduce both the complexity of the analysis and the costs of data processing, a subset of the data set was assembled including weight-health observations on the children from only six time periods. These periods were six months apart, starting in January of 1975 and concluding with June 1977—the month of the survey. This subset of the data set was passed to the MIDAS statistical system for analysis.

Kottar - Analysis

An unusual and distinctive characteristic of the KOTTAR data set is that the longitudinal component was constructed retrospectively at a single point in time. Every weight-health observation was taken on a child participating in the program in July and August of 1977. Therefore, the phenomena of drop-outs, "graduation," and, to some extent, entry at a time

in life later than 3 months, do not occur in this data set. A comparison of a set of children at two points in time involves only those children still participating in July and August 1977 and, in fact, except for those few children missed at either point of comparison, involves the same children at different stages of their lives.

One such comparison, a comparison of the six time periods included in the data subset used for analysis, shows a marked improvement of the study population over time—both in the aggregate and by age group. The percentage of malnourished children summed over all ages dropped from 50.4 percent to 45.5 where malnourishment was defined to be a weight less than 75 percent of the NCHS-CDC weight-for-age standard. The improvement is even more marked if the local standard and classification is applied—50.2 percent to 42.5 percent. This particular standard called for less weight gain among older children. This, coupled with the increased number of older children in later time periods, accentuates the improvement in the nutritional status of the population.

Several potential competing explanations of this improvement were tested. Two explanations--that the observed improvement reflects the often encountered natural recovery associated with aging or that it reflects the better status of newcomers to the program--were shown to be inadequate. However, a third competitor, the effects of a drought, was more difficult to analyze. Using this data set, it was not possible to separate the improvement due to recovery from the drought from the improvement due to program effect.

Finally, a multivariate analysis was performed to ascertain the determinants of malnutrition. Regression analysis was performed to "explain" the variance in July 1977 nutritional status with assorted socioeconomic variables. Of most interest in this analysis was the lack of explanatory power in the socioeconomic variables often thought to determine nutritional status. Two regression models were formulated. The first incorporated the variables of age, the square of age, the percent of standard at first observation, the ratio of observations with reported illness to the total number of observations, the number of children in the family,

and the sex of the child. The second model added the variables of income, education of household head, total land available, amount of family debt, and village type (coastal or inland). The R-SQR (coefficient of determination) for the first model was .25 with all explanatory variables significant. The R-SQR for the second model was .26 with only three (income, land, and vilage type) of the added variables significant. This illustrates how the village selection process narrows the range of socio-economic variables so much that individual differences within that range became relatively unimportant.

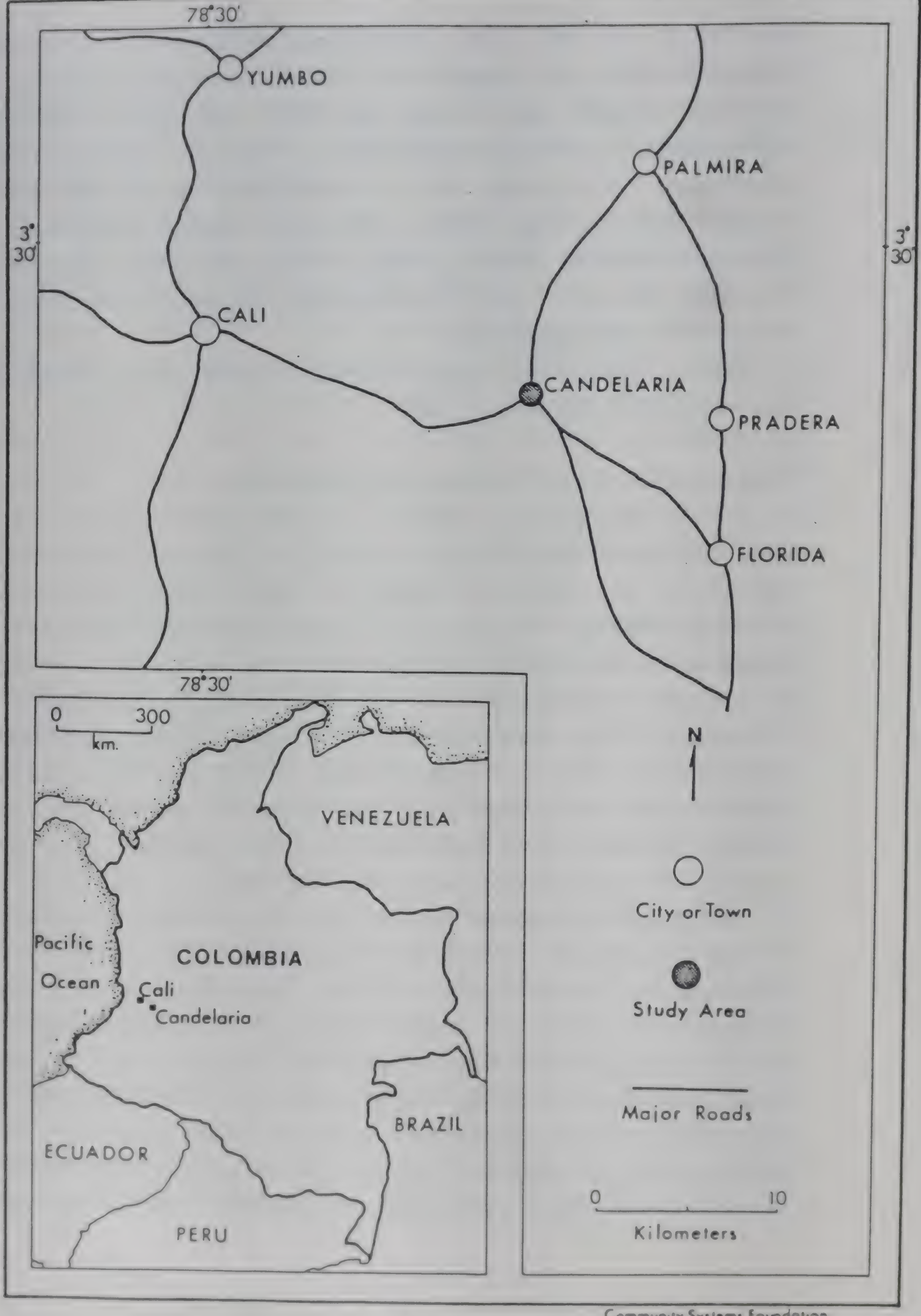
Volume III of this series is the complete study concerning the intervention and outcomes to date.

CANDELARIA AND CANDELARIA REVISITED

The city of Candelaria, Colombia, has had health-related interventions since 1958, when the Board of the Faculty of Medicine at the Universidad del Valle in Cali decided to establish a rural health center there for teaching purposes (Map 4). The Promotora Program (1968-1974) which generated the data used in our project is a step in this longer "life-cycle," an addition to an ongoing set of activities initiated by the faculty of the University. Forerunners of the Promotora Program included the establishment of a rural health center in 1958, a nutrition recuperation center in 1962 (which existed for two years), an outpatient treatment program including food supplementation (which replaced the recuperation center), and municipal sewer and water programs.

The Promotora Program represents an effort by the medical community in Cali to reach out into the homes of rural residents, rather than treat only those who visited a medical facility. When it began, the presence of the medical contingent was already accepted in Candelaria; there remained the embedding of the new program into the community. The strategy adopted was the training and use of volunteers from Candelaria as the home visitors. The tasks of the volunteers included provision of nutrition education, hygiene education, education on the utilization of other existing health services, data gathering, and referrals to the appropriate health

MAP 4 - PROMOTORA PROGRAM - CANDELARIA, COLOMBIA



service facility. This direct participation of community members increased the program's acceptability, while at the same time freeing the more highly skilled health care personnel for tasks more appropriate for those skills. The program remained viable for six years. In 1974, the government took over the staffing of the health center in Candelaria and the Promotora visits were phased out.

Candelaria - Data

The data in Candelaria consist of time series data on the anthropometric measurements, some morbidity information, and a rich collection of family socioeconomic variables. Eighteen hundred families participated at some time during the six-year period of the program, providing a total of 9,800 observations on 1,051 children.

In what to us is a unique activity in the nutrition field, the population of Candelaria was resurveyed in 1976, two years following program termination. The resurvey was administered to participants and nonparticipants in the program (most nonparticipants were new families and in-migrants). It included the gathering of anthropometric and socioeconomic data.

Candelaria - Analysis

During the course of the intervention, the malnourishment rate in the population dropped from 26 percent at the start to a stable figure of about 21 percent during the last three years of the program. This 20 percent decrease was found to be attributable to the program. Other conclusions from this initial study were as follows:

-- A strong relationship between the presence of diarrhea and malnourishment is shown both graphically and statistically. More important is the finding that the Promotora Program reduced incidence of diarrhea significantly (by 20 percent to 50 percent depending upon the age of the child and the duration of time the family was in the program).

-- A particularly surprising result was that the nutritional

status of girls improved far more than boys during the seven-year period. Based on the Gomez standard used during the period of the program, malnourishment in girls decreased from 36.1% to 24.8%, while the rate for boys went from 19.7% to 17.3%. That is, girl malnourishment dropped 31% and boy malnourishment dropped only 12%. Although the utilization of a sex-differentiated standard would change the relative position of boys to girls, there would be no change in the finding that girls improved far more than boys.

-- The education level of the parent, expenditures on food, and monthly income all affected the level of child malnourishment.

-- During the seven-year period, 1968-1974, there was a significant erosion in family purchasing power for the entire community. Although the average family income increased from 683 pesos to 1,251 pesos, real income, when adjusted for inflation, dropped to 566 pesos in constant terms. This resulted in a decrease in real expenditures on food of 8.7% over the seven-year period.

A function describing the relationship between infectious disease and food consumption was formulated. This relationship is important because it could lead to helping define what the best mixture of community level interventions should be.

-- Consumption of food also dropped, which increased the gap between need and consumption from 22.7% to 33% for proteins and 19.7% to 29.7% for calories. In addition, shifts occurred from animal proteins to lower priced vegetable proteins. Whereas the mixture was 47% animal and 53% vegetable in 1968, it was 37% animal and 63% vegetable by 1974.

-- The use of birth control more than doubled during the

Promotora Program. Families using some form increased from 18.7% in 1968 to 43.6% in 1974, with the largest change in use being the IUD. Birth order was shown to affect sibling malnourishment rate in Candelaria. It is therefore reasonable to presume that as family planning increases, there will be further positive effects upon the child malnourishment rate.

Differences between participants and nonparticipants two years after termination of the Promotora Program can be summarized as follows:

1. The nutritional status of children born into participating families following termination of the Program appears to be no better than children in nonparticipating families.

2. The nutritional status of children who actually participated in the program two years earlier is less clear. Some of our analytic indicators show an improved condition and others do not.

3. Children from families that participated two years previously appear to suffer less from diarrhea than children in nonparticipating families. Different methods of analysis generated inconsistent conclusions in this instance. However, since the mother's recall of diarrhea is a notoriously weak method of measuring prevalence and severity, underreporting is common. On balance, we are confident that this conclusion is valid.

4. Two years following termination of the Promotora Program, other diseases do not appear to be significantly affected by past participation, although this lack of relationship may be due to field-level difficulties in accurate measurement.

5. Fewer participating mothers in all age categories were pregnant than were their nonparticipating counterparts. Depending on the measure used, the pregnancy rate was roughly thirty percent lower in participants two years following termination of the Program. This finding holds true regardless of the analytics used and is probably the strongest relationship in the study. Since there was a government-operated family planning program in existence during the entire period, there is strong evidence that reinforcement arising from exposure to the outreach aspects of the Promotora Program resulted in considerable

synergism. The question of the level of intensity of effort necessary to have a meaningful impact on the target population is undoubtedly one of the most significant policy issues facing sponsors of interventions.

The initial evaluation and the revisitation study are presented in complete form in Volume IV of this series.

PRIMOPS

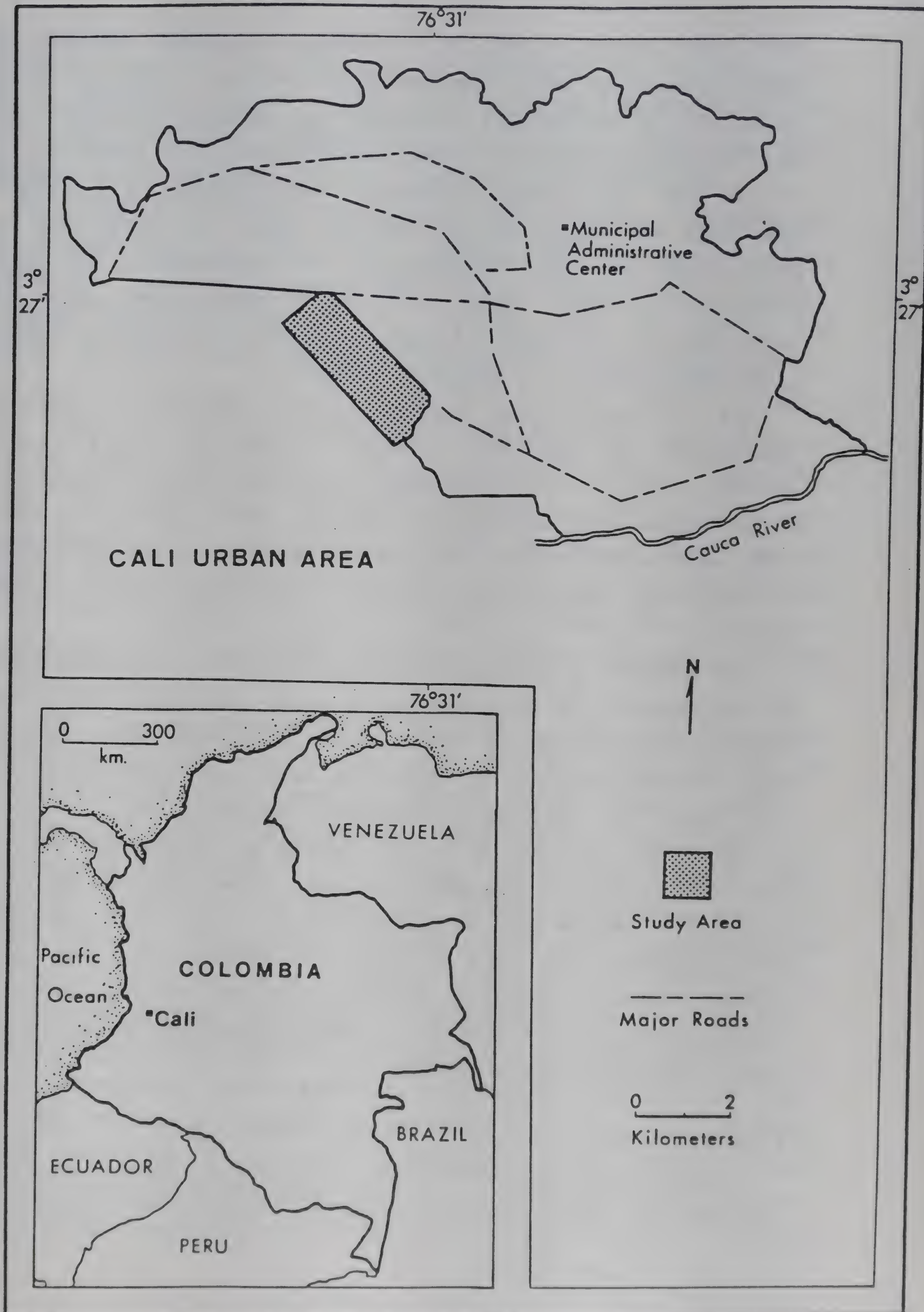
The Programa de Investigacion en Modelos de Prestacion de Servicios de Salud (PRIMOPS) is a joint effort by the Department of Health of the City of Cali and Department of Community Health of the Universidad del Valle in Cali, Colombia (Map 5). It represents an effort spearheaded by Dr. Alfredo Aguirre to "scale-up" the Candelaria program in an urban barrio of Colombia: La Union in Cali.

With funds provided by US/AID, Dr. Aguirre, along with his colleagues from the Universidad del Valle and public health researchers from Tulane University, designed an intricate training program for the Promotoras. The planning documents for PRIMOPS as a whole and the training program in particular suggest that it is one of the most carefully thought-out intervention of any we have reviewed. The guiding principle and belief of these entrepreneurs was that certain primary health care and diagnosis functions can be done by people recruited from the local community if they receive proper instruction and supervision. Moreover, both training and field operations can be accomplished at costs within national health planning budgets. The PRIMOPS program convincingly demonstrated that these goals can be accomplished in a poor urban neighborhood.

Although the project designs called for an assessment of the effect of the program on community health and nutrition, this question of attribution received less attention, and subsequent evidence suffers from the same shortcomings found in all the projects: competing explanations for changes observed are at least as likely as attributing the changes to the program actions.

Community embedding was achieved by recruiting promotoras through local **juntas** (neighborhood political groups) or health clubs and by calling

MAP 5 - PRIMOPS PROGRAM - CALI, COLOMBIA



Community Systems Foundation

upon these groups to help in establishing the physical facilities for the health posts out of which the promotoras worked. Also an important factor in the communuity embedding and subsequent durability was the close coordination with the already existing health care hierarchy of the City of Cali. The Program Co-director was an official of the city health department and the promotoras, auxiliary nurses and supervising nurses were, from the start, employees of that department. This made the transition to permanent status in the city's health system easier at the time when the University funding ended and the University researchers turned their attention to scaling-up problems.

The first step in this program was to complete a census of all households in the study area in order to identify eligible women and children. This was to be followed by periodic visits to households with eligible people every three months starting in 1973. Actually the rounds of visits were never completed that quickly (with five months the average and with one round taking ten months), due, in part, to a special immunization program.

The functions of the promotora were to provide elementary health and hygiene instruction in the home, and to identify and refer sick or high-risk children and/or pregnant women to the appropriate health service unit. The promotoras weighed the children to assess nutritional status and to refer second- and third-degree malnourished to special help.

Midway into the program, in 1975, funding was stopped and the program went into a period of complete inactivity. It was resumed again in 1976, and in 1979, became fully incorporated in the city health care system. The Universidad del Valle people, in an effort to keep their activities flowing, proposed a scaled-up program to cover other populations and regions with a similar service. These programs are continuing in 1980 under various funding arrangements.

PRIMOPS - Data

Several external evaluations were built into the PRIMOPS program from the outset and an internal data collection activity was established and

maintained. Because of the hiatus and the independent flow of funds to the external evaluators (a flow which called for final reports during the hiatus), the only study completed as of the summer of 1980 was of the efficiency of project teams in terms of cost of services delivered.

The internal data system (the anthropometric, health and immunization data on individuals) was never used. Unfortunately, the early data (pre-hiatus) was recorded in an entirely different format using a different individual identification scheme than the later data. As part of our project, we have spent long hours cleaning the data beginning with the 1976 rounds and, at this time, have two and one-half rounds in usable form. The data has many cases but few variables. For each promotora visit, the child's weight, age, diarrhea status, general health, immunization activity and use of medical facilities is recorded; however, for all the data except weight and age, the inability to distinguish missing data from the response "none," coupled with apparent inconsistencies by promotoras in coding responses, weakens our confidence in these other variables considerably. As a consequence we did not undertake multivariate analysis of the PRIMOPS data.

In the two and one-half rounds for which we have machine readable data, 12,770 children participated. The first and second rounds had enrolled about 8,000 children with about 3,000 new children enrolled in Round Two (an apparent 35% turnover). Only 10% of these are accounted for by control codes in the record of visits. The following categories were identified: moved, over-age (greater than 5 years old), refused interview, or died. The 25% unaccountable loss may be the result of high migration rates in the barrio but also, to some unknown degree, by errors in ID number recorded, round number or new/continue code entry.

Our efforts to correct cold PRIMOPS data far removed from the instant of collection led to several of our conclusions regarding the value of a rapid information feedback system which is reported upon elsewhere in this report. Out of a total of 21,916 records for the twelve thousand children we created a final clean data set of 20,655 observations, in which we accepted a 5.8% loss due to detected but unresolved errors. If we had

simply dropped each record containing a detected error, the data set would have been reduced by 4,303 observations (19.6%). Of course, correcting errors after-the-fact and retaining the record in the file entailed making certain assumptions which may or may not have been valid. These problems are discussed in greater detail in Volume V and Volume VII of this series.

The value of a program with records of periodic visits is the opportunity to obtain longitudinal data on individual children. PRIMOPS proved disappointing in this regard. The number of children with multiple observations dropped off rapidly from 6,639 with single observation, 4,377 with two observations and 1,754 with three observations, that is, 14 percent of all children were observed three times.

However, these numbers were large enough to generate characteristic curves (graphs displaying the percentage of children malnourished by age) with stable values by annual age groupings. The curves displayed the expected pattern of rising malnourishment peaking at 24-30 months and declining thereafter. Furthermore the percent malnourished decreased steadily through rounds one, two and three in about all age groupings. The figures for all ages are:

Round One	31.7%
Round Two	29.0%
Round Three	25.5%

This result, which is probably quite real, is difficult to attribute to the action of the Promotora Program alone. Additional analysis reveals that children entering the program in Rounds Two and Three are generally in better nutritional status than those continuing in the program. Those continuing in the program also show steady improvement from round to round but to a lesser degree than indicated by the round totals. An alternative explanation is the improved sanitary conditions in the barrios which resulted from continuously upgrading the quality of individual houses and the extension of public services to greater and greater proportions of the households; the services included drainage ditches, roadways, water, sewer and power connections.

The longitudinal data on individual children permitted construction of transition matrices showing changes in nutritional status from round to round. These tables indicate steady improvement of nutritional status in the community as more people shift from severe to mild malnourishment to normal status than in the reverse order. However, the tables also indicate that a higher proportion of severely malnourished transfer to **lost record** category than **mild** and **normal** groups. As the total in **unaccounted lost** is quite large, our confidence in the demonstrated improvement is weakened. Many problem families may have simply drifted out of the program.

Our major conclusion from the analysis of PRIMOPS is that even a well thought-out and clearly operational health care delivery program needs a theory-oriented data management system integrated into daily operations in order to assure survival of information useful for detecting a response to a changing condition whether planned or part of the changing environment of the program.

Volume V of this series reports details on the results of this study.

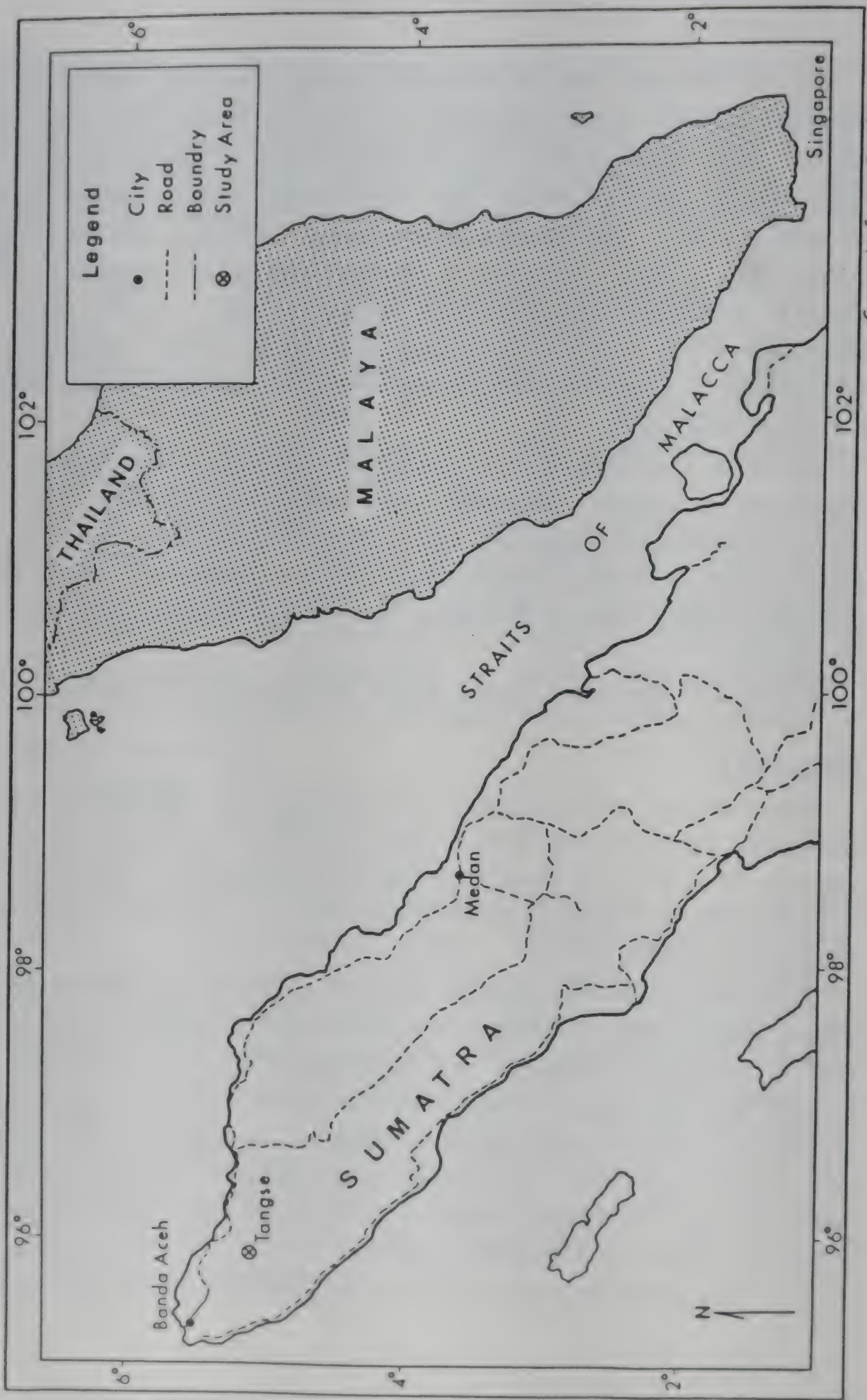
INDONESIA

The Indonesia intervention is run by Save The Children (SAVE) in Kecamatan Tangse in the Special Territory of Aceh on the island of Sumatra (Map 6). The Save The Children involvement in Aceh began in January, 1977, with funding from US/AID for a five-year Operational Program Grant (OPG).

In general, Save The Children attempts to play a role as catalyst to general development by taking a low profile over a sustained period of time in an effort to organize local people and foster appropriate linkages between the local power structure and the outside world. In theory, the process calls for the identification of problems and suggestion of solutions by the local community. In practice, SAVE mediates between the local community and the "outside world" in an effort to come up with realistic "fundable" options for development activity.

The Indonesian intervention is nearing the end of its community

MAP 6 - TANGSE PROGRAM - TANGSE, INDONESIA



embedding stage. That stage was initiated by a request from the governor of the special province of Aceh for non-governmental assistance in development work. As indicated by its name, Aceh Province is indeed special; it has a strongly independent population, governed by Muslim rather than civil law, and has never been integrated into the regular governmental structure of Indonesia. The process of gaining acceptance by the people of Aceh (not to be confused with the government in Aceh) has been a long and difficult task for the SAVE staff. The process used has been the provision of services (especially medical services) that were either not available in the study area or were available but not easily accessible to the local people.

SAVE, following its Community Based Integrated Rural Development (CBIRD) model, formed a community development committee consisting of representatives of eleven target villages. However, it became apparent that a separate committee was needed in each village. These were established in early 1978.

The projects which have been implemented at some level in Tangse are:

1. the establishment of village health posts staffed by trained volunteers from the villages;
2. child weighing programs with weight charts kept by the mother;
3. food preparation demonstrations;
4. soybean processing;
5. the construction of latrines;
6. the planting of home gardens;
7. the creation of fish ponds;
8. immunization of preschool and school-age children.

The degree of implementation of each program component varies from village to village. These are not "blanket" activities provided to all. Progress in community embedding has been retarded by military harassment

of SAVE staff, floods, and cholera outbreaks, and the "usual" difficulties in getting a community to accept a change suggested by outsiders.

Indonesia - Data

Save The Children initiated a data-gathering operation in part in response to the urgings of William Drake, principal investigator on the CSF project to analyze community-level nutrition interventions. During a visit to Aceh, Dr. Drake illustrated the use of weighing for diagnostic purposes. The SAVE staff put the technique into operation.

The number of data elements recorded is relatively few: the names of both mother and child, the child's age at the time of his/her first weighing, the sex of the child, and all weights taken for that child. The first set of data on 589 children consisted of multiple observations on those children, spanning a period of 18 months. The observations were not regularly timed and not distributed evenly in rate of occurrence over the 589 children.

Indonesia - Analysis

Analysis of so limited a data set is difficult. A comparison of the nutritional status of all children as of their first weighing to the status of those same children as of their last weighing indicates that there has been some improvement. (One must bear in mind that the time difference and number of observations between first and last observation is different for each child.) Furthermore, it can be shown that the most malnourished showed the greatest improvement. However, this phenomenon can be explained partially as an improvement resulting from the aging of children, so that more were past the age of maximum risk as of their last observation.

Volume VI of this series reports the results of this study in more complete detail.

HONDURAS

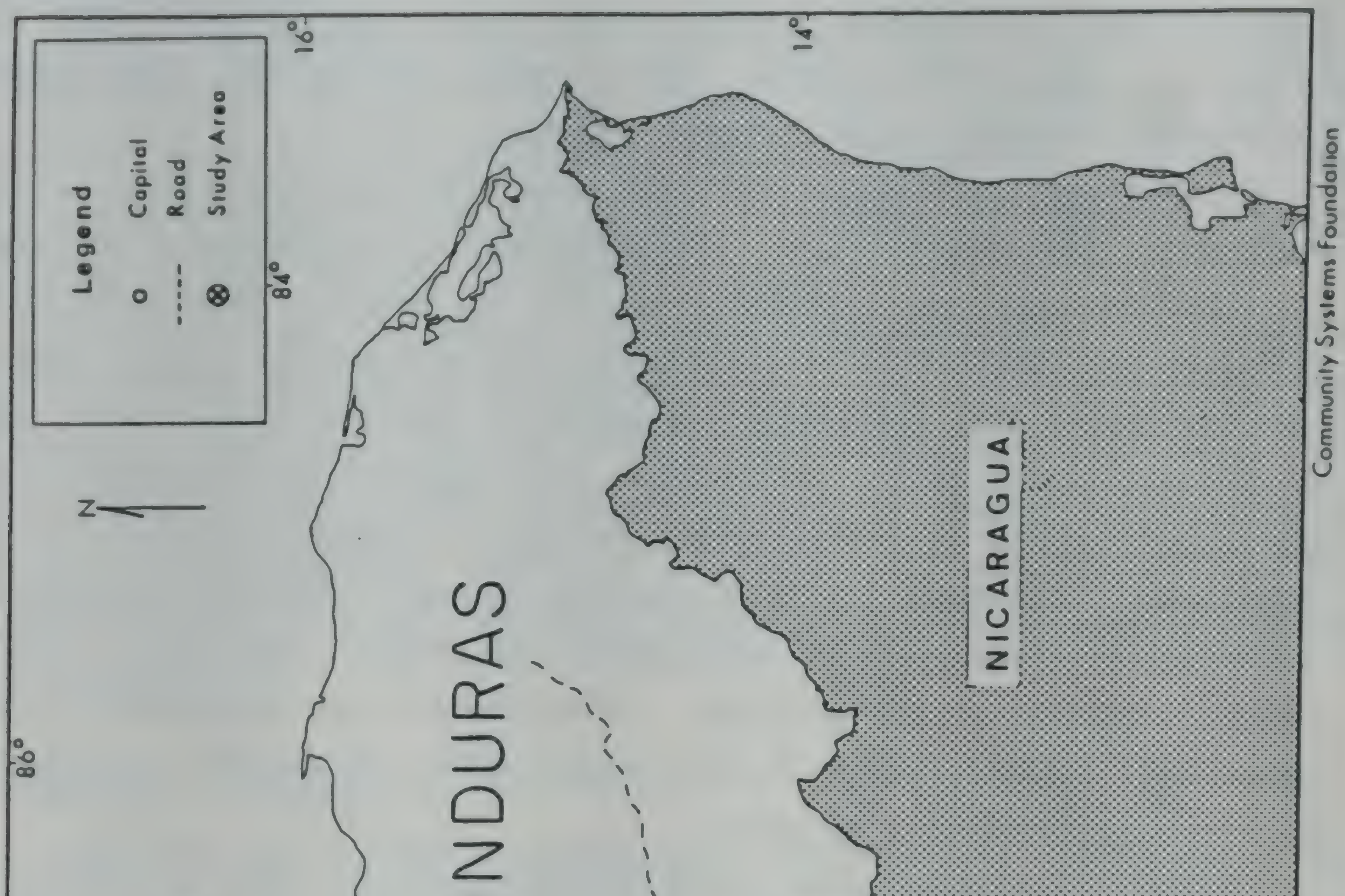
The Honduras intervention is run by Save The Children (SAVE) in the Pespire region of Honduras (Map 7). The government of Honduras, recognizing the problems existing in the Pespire area, invited Save The Children to work in the region. This program followed the same general development strategy as the Indonesia program.

The Honduras intervention ran into some very real difficulties in the community embedding stage. Recognizing the seriousness of the food shortage resulting from prolonged drought in Pespire, the SAVE staff undertook the distribution of food supplements provided by CARE. Unlike all other CARE food distribution programs in Honduras, this program was of the take-home variety. SAVE made food available to all comers. The increasing demand as news spread caused a monthly shortfall in the commodity. The result was that the food became a "political" resource in the hands of the community development committee. The result was a near breakdown in the program. The situation was aggravated by the government medical officer for the region, whose directives kept SAVE from serving those most in need.

Despite these problems, SAVE has continued its work in Pespire—at reduced pace. The food program was stopped in January, 1978. Other components of the program include:

1. construction of a piped water system;
2. construction and maintenance of fish ponds;
3. an education program on the use of already existing latrines;
4. construction of new latrines;
5. formation of a cooperative to market processed food products made from locally available foods.

<u>MAP 7 - HONDURAS PROGRAM - PESPIRE, HONDURAS</u>



Honduras - Data

In response to the CARE requirements for determining food allocations, SAVE initiated a weighing program in the region. At first, weights for each child were both recorded and plotted monthly on a growth chart. Simultaneously, the lists needed by CARE to set allocations were compiled. A visit by the SAVE home office (including several consultants—one from our own project) led to a revision of the data collection procedure. A new simplified form was designed which led to the recording of height and "current" health status as well as weight. Children were brought to the health center every three months rather than monthly, as part of the revision.

The original forms, as well as the first round of data collected with the revised forms, were brought to Ann Arbor to facilitate analysis. These were keypunched and checked for errors. The resulting data set was placed into the MICRO data base management system. As noted, the food program was stopped in January of 1978 and, with it, the weighing program. The data collected between the summer of 1977 and the start of 1978 were not added to the data set.

The final data set consists of weights taken in three-month intervals for 805 children living in five villages. There are five rounds of data spanning a time period of one year. In one village only, a sixth round of data reached Ann Arbor.

Honduras - Analysis

The only type of analysis that could be done with this data set was the tracking of community nutritional status as measured by the ratio of a child's actual weight to a weight-for-age standard. At first, community nutritional status appears to improve dramatically over time. However, the number of children participating increases steadily over time. In the fourth and fifth rounds, newcomers to the program are markedly better off than those who had begun participating earlier. This creates the appearance of positive program impact in the aggregated data; however, the original participants show no improvement above and beyond that which

could be attributed to aging. (The new additions to the program were healthier due to a policy set by the local government doctor to stop serving the most seriously malnourished.)

Volume VI reports this study in more complete detail.

THAILAND

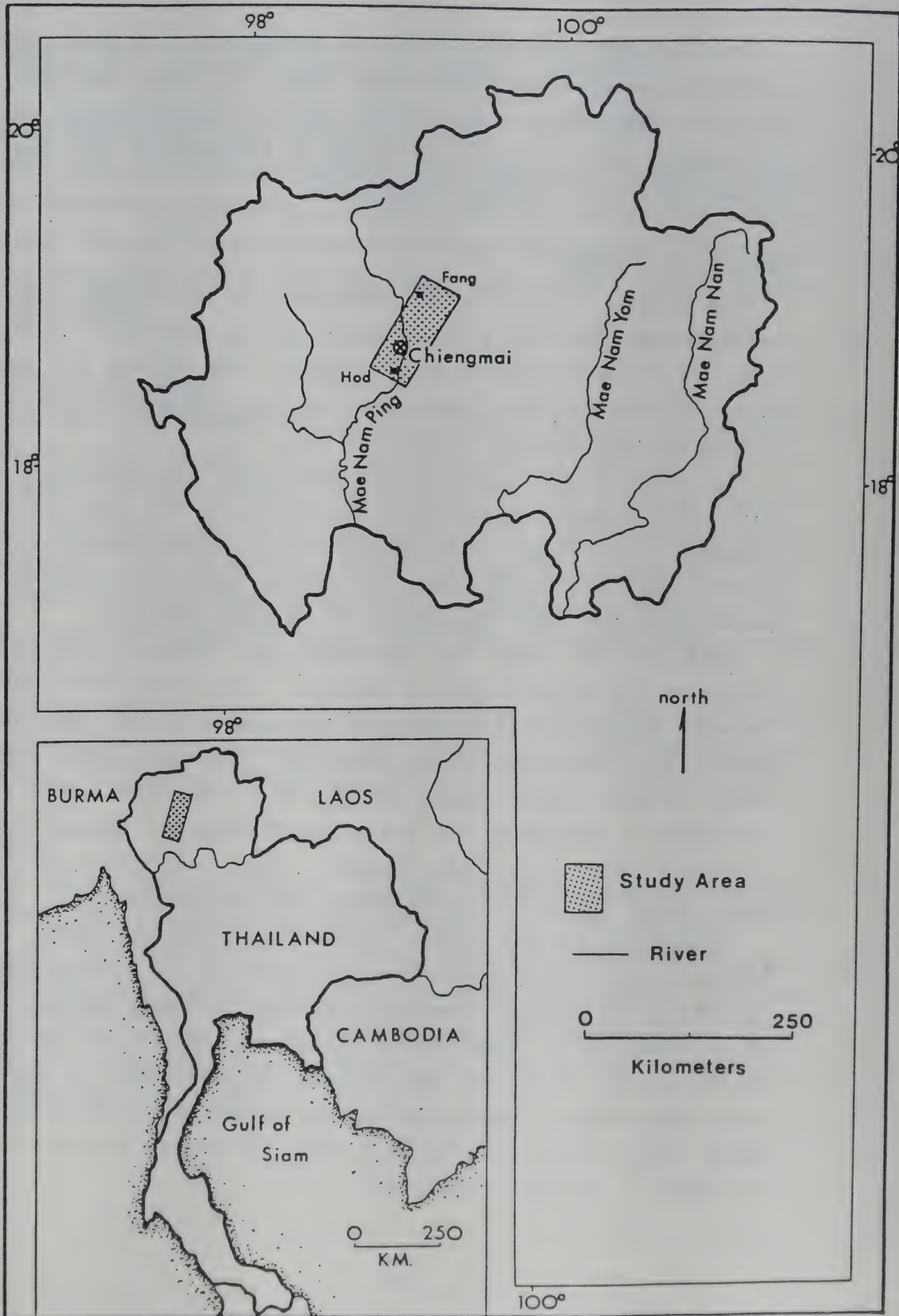
The Thailand Rice Fortification Project, conducted by Drs. Stanley N. Gershoff and Robert D. McGandy of the Harvard School of Public Health, was more of a research study than an intervention—a study to measure the health benefits to be derived from the fortification of rice with lysine, threonine, thiamin, riboflavin, vitamin A, and iron. At the time of the project's conception (1968), scientists were arguing the issue of the wisdom and efficacy of large-scale protein fortification. This was a field study to provide more evidence for that debate.

The Chiang Mai region of Thailand was selected because the study area had to: (1) have a population that derived the bulk of its calories from rice; (2) provide evidence of protein malnutrition; and (3) have infectious disease not so rampant as to obviate the beneficial effects of the program (Map 8). Although the main component of the program was rice fortification, day-care centers were constructed in some villages to demonstrate the good intentions of the intervenors, and a small-scale education program was conducted in the villages to secure the cooperation of the villagers. These last components provided the community embedding components for the study. The study was terminated in January of 1975 (it was actually begun in 1971) without any scaling-up.

Thailand - Data

Children from the 29 villages were given physical examinations twice a year—once in the dry season after harvest and once in the monsoon season. Anthropometrics plus some biochemical tests were made, and clinical observations were systematically recorded. Hand x-rays were taken yearly and morbidity data were collected starting in 1972. (The data are the most extensive medical data available to CSF.)

MAP 8 - RICE FORTIFICATION PROGRAM - CHIANG MAI, THIALAND



Community Systems Foundation

The 29 villages were divided into five groups, with each group receiving a different combination of fortification and/or day care centers. One group was designated as a pure control, one as a placebo control with day care centers, two with different types of fortification and day care centers, and the fifth with just fortification.

Thailand - Analysis

The analysis reported in the literature by the original study team concluded that there was a major change in the population during the study but no positive effects of fortification. The children who received little or no fortified grain experienced the same change as those who received the grain regularly (Gershoff, 1977).

The conclusion was reached by the study team after examining mean heights and weights of participating children at the start and at the conclusion of the study. Comparison with middle-class Bangkok children provided evidence that the Chiang Mai children remained deficient in these measurements even after experiencing the secular change.

Using the same data, we reproduced this analysis comparing the children to the Harvard growth standard. The results were similar; however, looking only at performance with respect to the weight-for-age measure, the improvement almost disappears. Whereas heights increase during the study period, weights increase at a lesser rate. As a result, malnutrition as defined by the Gomez classification and Harvard standard diminishes only a very small amount. Again, the control and placebo control groups fare as well as the groups receiving fortified grain.

Explanations for this observed outcome are purely speculative. First, the chosen fortification proteins may not have been the limiting factor in the diets of the children. Second, the children of Chiang Mai may simply not eat enough—even though food is available—because of the hot climate and the boredom of the rice diet. Finally, the determinants of nutritional status and growth in Thailand may be much broader than either quantity or quality of rice intake, and changes in those determinants may have masked the effects of fortification completely.

Volume VI of this series reports the findings of this effort.

CROSS-SECTIONAL AND OTHER STUDIES

The seven interventions described above generated the eight principal data sets used in this project. However, use was made of two cross-sectional data sets to test methodological questions, and the experience of Barton Burkhalter of CSF with the Papago Indians had a major influence on our thinking.

Dominican Republic

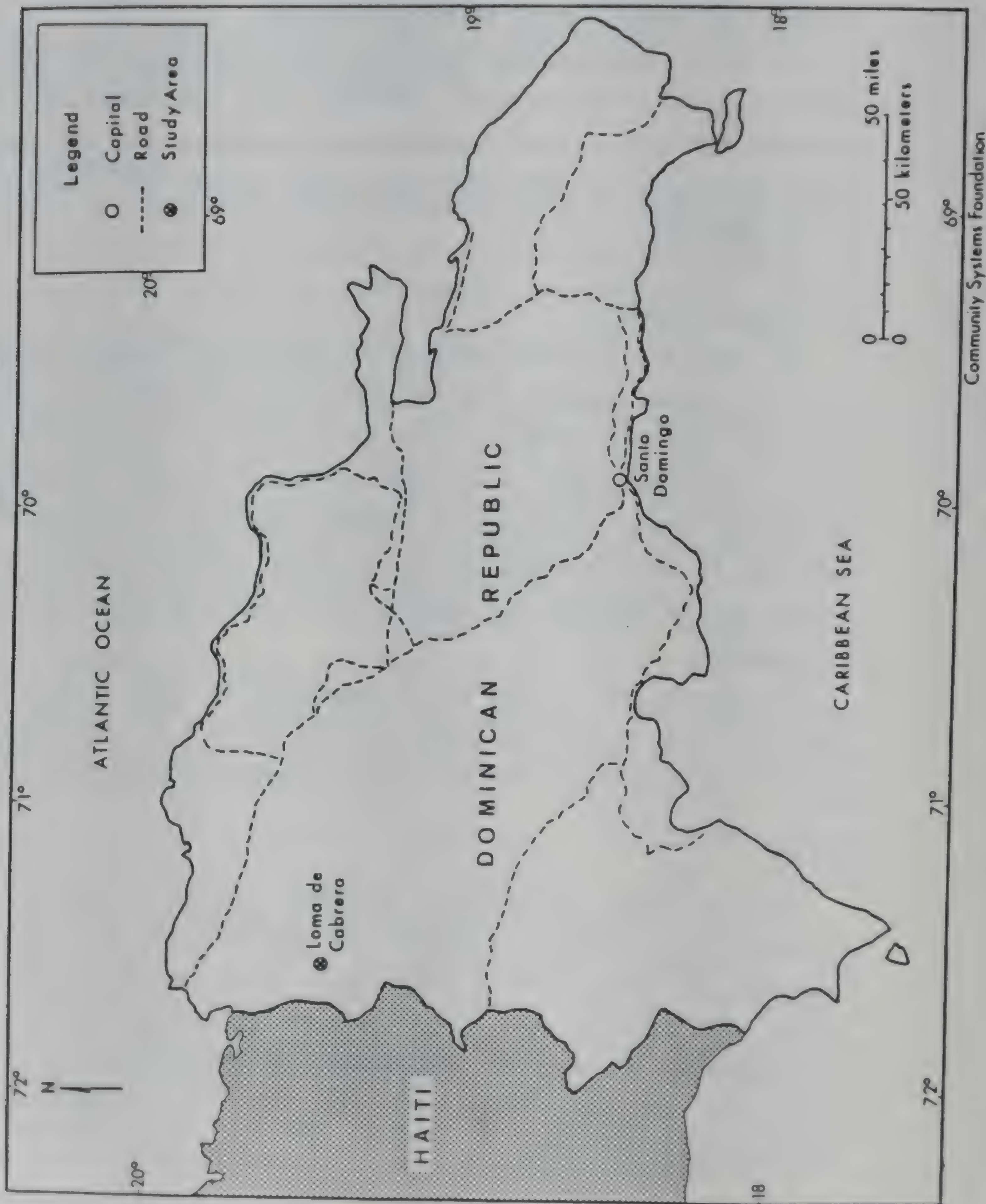
The Save The Children staff in the Dominican Republic administered a survey to obtain baseline data on the Loma de Cabrera region in the Dominican Republic (Map 9). The SAVE staff in the Dominican Republic was in the early stages of its implementation of the Community Based Integrated Rural Development (CBIRD) approach to development. Their hope was to use the results of the survey to identify problem areas in the region, and to establish the baseline from which subsequent progress could be measured.

The survey was coded and keypunched at CSF, and a preliminary analysis of the data was returned to SAVE staff in the Dominican Republic and the home office in Connecticut (Community Systems Foundation, 1978).

CARE

In 1976, CARE evaluated its preschool take-home feeding programs in Colombia, the Dominican Republic, and Pakistan, and its on-site feeding programs for preschoolers in Tamil Nadu, India, and Costa Rica (Anderson, 1977). The data included socioeconomic data, consumption data, and anthropometrics for both participants and nonparticipants in the CARE programs. The data have been used by CSF to test some of its methodological hypotheses.

MAP 9 - LOMA DE CABRERA PROGRAM - DOMINICAN REPUBLIC



Papago Indians

For the past year, Barton Burkhalter of the CSF staff has been observing and evaluating a variety of the interventions in progress for the Papago Indians in Arizona. While numeric analysis of these interventions was not formally part of this research project, the insights obtained from a review of the nutrition-related programs and the information system utilized by the Indian Health Service has been very helpful.

LEARNING FROM THE ANALYSES

The data sets and analyses described in this chapter form the "data" for the larger project to analyze community-level nutrition programs. Although much was learned about the separate interventions (see Volumes II through VI), this final report must generalize from that knowledge.

An immediately apparent feature of the synopses of the separate analyses of the interventions is the varying degree of confidence held in the findings. It is highly unusual for a research team to acknowledge a range of confidence in this way. Typically, analysts of social programs are tempted to create an aura of "high science" in presenting their findings. Although results are rarely conclusive, they are often presented with dogmatic assurance. However, to be realistic, analyses of social systems are quite subjective and are highly dependent on the judgments and assumptions of the analyst. It is our belief that the reader should be apprised of those judgments and assumptions and must develop his/her own estimate of the plausibility of the results.

Table 1 summarizes our own confidence in the findings with regard to the impact of the seven interventions. Our level of confidence is the subjective sum of our separate confidence in the quality and depth of the data, the consistency of our statistical analyses and, to some degree, our intuitive feeling toward the subjective and/or non-quantifiable dimensions of the intervention and its setting.

In accordance with the research strategy described in the introductory chapter, we have taken care to present competing explanations of the observed outcomes and to state which of those possible explanations is

TABLE 2 - 1

SYNOPSIS OF CHARACTERISTICS OF SELECTED PROGRAMS USING PRIMARY DATA

	Name & Location	Approximate Population Observed & Period of Data	Observed Outcome	Possible Attributions or Explanations for Outcome	Most Likely Attribution(s)	Confidence of Attribution	Type of Control Utilized
1	Candelaria, Colombia	1200 families (1968-1974)	• Malnourishment dropped from 29% to 21% (25% reduction) • Diarrhea prevalence decreased	• Program effects • Improvement in environment • Reduction in disposable income • Aging population • Target population change • Selection biases	Promotora Program effects	Strong	Reflexive, statistical
2	Candelaria, Colombia (Revisited)	500 families (1976)	• Reduced diarrheal disease • 30% reduction in pregnancy rate	• Program effects • Target population change • Change in environment	Synergism between Promotora Program and other programs	Strong	Participants & nonparticipant groups (statistical)
3	Kottar Social Services Society Tamil Nadu, India	4,000 families in 21 villages (1975-1977)	• Malnourishment dropped from 50.2% to 42.5% (15.3% reduction)	• Program effects • Environmental changes (drought) • Aging population • Selection biases	Program effects and/or improved environment	Strong	Reflexive, statistical
4	Esperanca Santarem, Brazil	460 children in 4 villages (1977-1979)	• Malnourishment dropped from 48% to 33% in program village (31% reduction)	• Program effects • Economic conditions improved • Environmental conditions • Aging of population	Program effects combined with improved economic conditions	Moderate	Treated & untreated communities, statistical
5	Prinops Cali, Colombia	11,700 children (1976-1977)	• Malnourishment dropped from 31% to 27% (13% reduction)	• Program effects • Change in target population • Improvement in water & sewer • Population aging	Program effects combined with improvements in environment and target population	Moderate	Reflexive, statistical
6	Rice Fortification Project Chang Mai, Thailand	1200 children in 29 villages (1971-1975)	• Slight reduction in malnourishment	• Economic improvement • Environmental changes • Program effects	Environmental changes	Moderate	Variously treated communities, reflexive, statistical
7	Community Development Project Tangse, Indonesia	500 children in 11 villages (1977-1979)	• Slight reduction in malnourishment	• Aging population • Program effects • Environmental conditions changed • Selection biases • Target population change	Aging population combined with program effects	Weak	Reflexive, statistical
8	Community Development Project Pespire, Honduras	500 children in 11 villages (1977-1978)	• Malnourishment dropped from 28.9% to 23.1% (20% reduction)	• Target population change • Selection bias • Drought condition changes • Program effects	Selection biases combined with possible changes in target population and environmental conditions	Moderate	Reflexive, statistical

most plausible. The confidence of attribution refers to the plausibility of the chosen explanations—not to the success of the intervention. We had strong confidence in our attribution three times and moderate confidence in four others.

We can now turn to the sources of eroding confidence, with special attention given to the ways that the research strategy can minimize the problems associated with those sources.

CHAPTER III

FINDINGS CONCERNING THE MEASUREMENT OF NUTRITIONAL STATUS

We have argued that the ultimate measure of success for a nutrition intervention is the long-term change in nutritional status of individuals and the community as a whole. Unfortunately, nutritional status is not easily determined and, for a given method of measuring nutritional status, the definition of malnourishment is a subject of debate. Even in the most modern clinical setting, medical experts will disagree over the nutritional status of a given individual; and in the best experimental field setting, nutrition specialists will argue over the nutritional status of the community. This chapter addresses the problem of measuring nutritional status and the implications of the selection of a method of measurement.

METHOD OF MEASURING NUTRITIONAL STATUS

Many articles in the nutrition literature describe the anthropometric and biochemical tests that can be used to assess the nutritional status of children under six years of age. We will summarize the advantages and disadvantages of each method in this section. Readers familiar with the literature may choose to skip this section, with the exception of the subsection entitled Anthropometry, which establishes a notational convention used in this report. For more detailed discussions, the reader is referred to articles by Keller, Donoso, and DeMaeyer (1976), Jelliffe (1966), and Zeitlin (1977).

A Basis for Selection

Several factors influence the choice of method used in a particular program to assess nutritional status. These include:

1. Reason for assessing nutritional status; for example, is it a

NUT100 3783

COMMUNITY HEALTH CELL LIBRARY AND DOCUMENTATION UNIT

single survey (observation) to determine the nature and extent of nutritional problems or gather baseline data, or is it part of a **continuous** or **periodic** assessment activity to screen individuals for health care or provide data for an early-warning surveillance system?

2. Specific cultural attitudes and the availability of relevant data; for example, the ease of obtaining accurate ages if using an age-specific measure.

3. The skill and training of the individuals who will be taking the measurements and computing the nutritional status.

4. The ease or difficulty of obtaining the required data; that is, special equipment required, special training, ease of transporting equipment, etc.

5. The location where data will be gathered. Will individuals be coming to a central location (health post, center, etc.), or will the survey team or health worker visit the dwelling?

6. Size of population to be surveyed, time available to do the survey, and financial constraints.

All of the above considerations are interrelated and have implications for the precision and reliability of the data collected, regardless of the method of nutritional assessment used.

With only a few exceptions where our opinion was requested, the measures actually used in the ten data sets comprising our data base were selected by the change agents in the field. In most instances, we had no choice but to use the measures that had already been selected. The anthropometric, biochemical, and clinical data extant in each of the data sets analyzed is summarized in Table 3-1.

Standards and Measures

As with most illnesses, malnutrrition manifests itself in individuals in varying degrees. The severity of a particular case can be determined by comparing the appropriate characteristics of the afflicted individual with those of a normal individual. One must select characteristics that most appropriately mirror nutrition and specify what is normal for those

TABLE 3-1

ASSESSMENT OF NUTRITIONAL STATUS
IN DATASETS STUDIED BY THIS PROJECT

Name of Program	Weight	Height	Arm Circumfr.	Head Circumfr.	Chest Circumfr.	Triceps Skinfold	Subscapular Skinfold	BIOCHEMICAL	CLINICAL
ESPERANCA	X	X*	X					Hematocrit	Clinical Assessment, i.e., skin, hair, etc. Stool Exam
CANDELARIA	X	X*							
CANDELARIA REVISITED	X	X	X	X					Clinical Assessment, i.e., skin, hair, etc.
PRIMOPS	X								
SAVE THE CHILDREN HONDURAS	X	X*							
SAVE THE CHILDREN INDONESIA	X								
THAILAND - RICE FORTIFICATION	X	X	X	X	X	X	X	Hemoglobin Hematocrit	Clinical Assessment, X-Ray - Hand, Stool Exam
KOTTAR INDIA	X								
DOMINICAN REPUBLIC	X								
CARE (5 datasets)	X	X	X						

* Height included in only 1 of the time-series observations

characteristics. The specification of normality for a particular characteristic is a standard.

Anthropometry. In preschool children, one characteristic that shows the ravages of poor nutrition is growth. Growth can be observed by looking at many of the physical characteristics of a child. Those most often used are weight and height. Unfortunately, these physical measurements are not constant for all preschool children; therefore, a standard must be formulated with respect to some other factor that serves as a basis of comparison—for example, age. Thus, a growth standard is a postulated relationship between the physical characteristic and the factor used to facilitate comparison.

In what follows, we will discuss the three ways to look at weight and height deficiencies most often used in practice. The first is the ratio of the observed weight of a child to the normal weight of a child of the same age—weight/weight(age). Deficiency with respect to this measure reflects either acute (short-term) malnourishment or chronic (long-term) malnourishment, or both. The second is the ratio of the observed height of a child to the normal height of a child of the same age--height/height(age). Because the structural deficiency leading to reduced height takes time to develop, this measure reflects a history of nutritional deficiency (chronic malnutrition). Finally, one can compare the observed weight of a child to the normal weight of a child of the same height—weight/weight(height). Because weight loss can be rather sudden, this ratio is an indication of the severity of acute malnourishment.

The assumption underlying the use of each of these ratios for monitoring growth performance is that the well-nourished child will grow in a normal fashion—that is, will follow a growth trajectory that approximates that of some predetermined standard. We know that a malnourished child will not grow normally; therefore, a significant deviation from the standard is an indication of malnourishment. The degree of deviation which constitutes a malnutrition problem is often expressed as a "classification" scheme--an indication of the magnitude of deficiency associated with

selected levels of risk.

The following paragraphs discuss the relative merits and drawbacks of these most often used measures of nutritional status.

Weight/Weight(Age). Weight is the most frequently collected anthropometric measurement. It reflects the total body mass of the child at a particular point in time. The advantages of using weight to assess nutritional status are that the data are relatively easy to gather in the field, and no clinically trained staff or specialized equipment are required. Apart from the problems of obtaining accurate weight, correct birthdates, and selecting a reference standard, the major disadvantage of using weight for age is that it is a composite indicator of both acute and chronic malnourishment. Thus, using weight for age can overstate the actual level of malnourishment in a particular population, because it will include children who are simply small for their age because of prior stunting or genetic differences.

If the approach of measuring changes in weight-for-age in the **same** children is used in a longitudinal investigation, more accuracy is required in the measurement procedures, because errors become amplified when comparing differences. For instance, suppose that a two-year-old weighed 10 kg. at the first measurement and 6 months later he weight 12 kg. The improvement in weight would be 2 kg. Yet if there was a 5% error in only one of the measurements, making the correct second weight 11.4 kg. instead of 12 kg., the change introduced in the difference between the 2 weights would be .6 kg., or 30%.

Height/Height(Age). Height is an indicator of the structural growth of a child. It is more difficult to measure accurately in a remote setting than weight and, therefore, is found more frequently in cross-sectional studies, where the difficulties of measurement must be overcome only once.

The advantages of using height-for-age to assess nutritional status are many. The distribution of height is narrow, so the effects of small

measurement errors are magnified. Height/height(age) is not a good indicator of nutritional status in infants or young children because of the time lag before an impact on skeletal growth occurs. The interpretation of height deficits and the degree of stunting are closely related to age. A normal child takes three to four years to double its length, so normal increased in length are slight in relation to length at birth. Genetic or ethnic differences in stature are also reflected in height measurement.

The equipment required to measure height is relatively simple, but cumbersome. Provision must be made for taking recumbent measurements on infants and young children, as well as standing measurements on older children. This does not present a major problem if measurements are done at a central location, for example, a health post, but is more difficult if equipment has to be transported from house to house.

Weight/Weight(Height). Using a weight/weight(height) ratio to assess nutritional status adjusts for the influence of height by comparing weight with a reference group of the same height, rather than of the same age. Weight/weight(height) is expressed as the ratio of the reference median weight for median height at each age. It is a measure of the thinness or fullness of the body, and indicates present nutritional status.

The major advantage of using a weight/weight(height) ratio for assessing nutritional status is that it is considered to be relatively age-independent. The disadvantages are that it requires two anthropometric measures—weight and height—and it does not provide data on the duration of malnutrition or of past malnutrition. Because of the difficulties of measuring height in infants, weight/weight(height) is a more appropriate method for assessing nutritional status in children age two or older.

Weight/weight(height) and height/height(age) can be combined to form a two-by-two grid that differentiates children into four categories: wasted and stunted, wasted but not stunted, not wasted but stunted, not wasted and not stunted (normal). This is called Waterlow grid. It provides the maximum amount of information regarding nutritional status that can be obtained from weight, height, and age (Waterlow, 1972).

The following paragraphs discuss some of the other anthropometric measurements that have been used by some intervenors in the past.

Arm Circumference. The arm circumference measurement is used to assess the muscle mass or protein reserves of the child's body. Muscle wasting is a general characteristic of protein-calorie malnutrition, so arm circumference can be used to determine nutritional status. When combined with the triceps skinfold measurement, the fat and muscle circumferences and cross-sectional areas can be computed for the upper arm, thus indicating both calorie and protein reserves.

Arm circumference has been widely touted as a screening and surveillance technique because it requires virtually no equipment, is easy to determine, and is not age-dependent between one and five years. It is adequate for determining the level of malnourishment of a community or specified population, but does not discriminate sufficiently to be used alone in determining the nutritional status of individual children for diagnostic purposes.

The measurement is made by locating the midpoint on the upper arm between the shoulder (acromion process of the scapula) and the elbow (olecranon process of the ulna). The midpoint of the arm must be accurately pinpointed, which requires a trained individual to palpate the shoulder and elbow to locate these processes. A flexible steel or fibre-glass tape is usually recommended for obtaining the measurement. It can be marked in centimeters or color-coded.

There is a relatively large standard error in measurement which, combined with the narrow distribution of data, necessitates large groups of children being measured to obtain an accurately scaled assessment of nutritional status. Errors generally result from the tape being placed incorrectly on the midpoint or being pulled too tightly. Thus, there are relatively high inter- and intra-examiner errors.

Thigh Circumference. The thigh circumference is a measure of physical growth and development. It can be used as a simple screening and surveillance technique, particularly for children under two years of age. The thigh circumference has many of the same advantages and disadvantages of the arm circumference measurement. Its use is most appropriate in areas where there are very limited health facilities, few scales, and a low literacy rate. Thigh circumference will indicate those severely malnourished children, but does not differentiate between normal and mildly to moderately malnourished children.

Thigh circumference is age-independent between twelve months and five years. The measurement requires a standard tape or color-coded tape, so the values can be recorded in centimeters or colors. Some of the disadvantages of thigh circumference are: (1) the measuring point on the mid-thigh must be located accurately; (2) modest older children can present problems; and (3) swelling due to edema causes severely malnourished children to appear normal in thigh circumference. Inter- and intra-observer reliability is also a source of error.

Chest/Head Circumference. The chest/head circumference ratio can be used to detect protein-calorie malnutrition in young children between six months and five years. It is not generally used as an individual screening technique but is most often used as a community indicator of the level of malnourishment in the child population.

The two circumferences are equal up to six months of age, after which the skull grows more slowly and the chest more rapidly. A ratio of less than one can indicate a failure to develop or a wasting of the muscle and fat of the chest wall. There is some evidence to suggest that head circumference alone can also reflect undernutrition that may have occurred in the child's past (chronic malnutrition), that is, before age two. The head circumference measurement can be used with a reference standard, though variability because of ethnic or genetic differences should not be overlooked.

The measurement chest/head circumference has the advantage of being

easy to conduct in the field. A flexible steel of fibre-glass tape can be used for obtaining both measurements. The chest measurement is more prone to error because of the problem of crying children and irregular breathing. If a tape is not available, the two circumferences can be compared using string.

Triceps Skinfold. The triceps skinfold measures the subcutaneous fat layer over the triceps muscle between the shoulder (acromion process) and the elbow (olecranon process). The triceps thickness estimates the contribution of subcutaneous fat to the arm circumference and gives an indication of the calorie reserves in the child's body. This method is most often used in conjunction with other nutritional status assessment techniques.

The measurement instrument is a skinfold caliper. This has rectangular jaws with a flexible head and exerts a constant pressure of 10 gm/sq. mm. The same midpoint on the upper arm that is used for the arm circumference measurement is used for the triceps skinfold; however, the measurement point is on the back of the arm over the triceps muscle. The skinfold is lifted up between finger and thumb and the calipers applied about 1 cm. below the examiner's fingers.

The triceps skinfold measurement is difficult to measure accurately. Most studies recommend that two or three measurements be made on each child and the results averaged. Thus, although requiring minimal equipment, the triceps skinfold is difficult to implement in the field, because it is subject to much measurement error and requires highly trained examiners.

Biochemical Tests. Several biochemical tests can be used as indices of malnutrition, but most are impractical to apply in a field setting in developing countries. In any case, biochemcial tests should always be used in conjunction with an anthropometric method for assessing nutritional status. The objective of most biochemical tests is to assess the nutrient levels in the body, as these levels may be affected before clinical and

anthropometric changes appear.

Nutritional anemia, vitamin A deficiency, and protein-calorie malnutrition are major problems in many developing countries. With the exception of hemoglobin and hematocrit, most biochemical tests are too costly and time consuming to perform in the field. The protocol for collecting and storing specimens also contributes to the difficulties.

Hemoglobin or hematocrit are the biochemical tests most frequently conducted in the field. The results are used to determine the prevalence of nutritional anemia. Hematocrit determines the volume of packed red corpuscles in the blood, while hemoglobin gives the red cell count. Both require only a finger-prick of blood, do not require refrigeration of the specimen, and are not affected by a recent meal. The hematocrit is easier to do in the field because it requires only a centrifuge which can be run from a generator or car battery and does not require highly trained lab technicians.

Hemoglobin levels can be measured by chemical analysis of iron, measurement of segments in the blood, and measurement of the physical properties of blood. Instruments and equipment required range from simple to very expensive.

In both the hematocrit and hemoglobin the greatest potential source of error is in the collection of the blood, the quality of the pipettes, and inaccurate interpretation of results. Like the anthropometric measure, there are suggested cut-off points defining hemoglobin and hematocrit deficiency. For example, WHO guidelines are that thirty-one to thirty-four percent scores on hematocrit indicate mild anemia, while anything lower is severe anemia.

Clinical Tests. Some programs that screen for malnutrition in a target population also include a clinical assessment of the physical appearance of the child. The clinical signs most related to protein-calorie malnutrition and vitamin A deficiency include muscle wasting, edema, changes in skin and hair, and Bitot's spots and night blindness. Generally, these signs appear in cases of severe malnutrition but are not useful for

detecting the earlier stages of malnutrition. While it is relatively easy to train someone to look for clinical signs in the skin, hair, eyes, mouth, abdomen, and musculo-skeletal system, it is always a subjective evaluation on the part of the observer.

IMPLICATIONS OF USING ANTHROPOMETRY TO MEASURE NUTRITIONAL STATUS

At the beginning of this chapter, we stated that the measures of nutritional status used throughout this study were selected by the change agents responsible for collecting the data in the field. Following common practice, these agents relied on weight, age, and in some cases, height data. In this section, we focus on some of the implications of this practice on analysis and interpretation. Although anthropometry is still the most efficient and most accurate way to quantify malnutrition in the field, it proves to be quite troublesome for the analyst concerned with evaluation.

Standard Selection Can Change Observed Outcomes

There is little agreement in the field on whether it is appropriate to use a local standard that accounts for genetic or cultural differences in populations or some absolute standard that all populations should achieve. We cannot resolve this problem; it is a subject for additional scientific study. We can point out that most "normal lines" are approximately parallel. Thus, an analytic methodology using the continuous underlying variable formed by computing the observation to a reference standard (not a categorical classification based on cut points) will yield approximately the same result for all "normal lines." We have computed all of our ratios in this study using the standard used at the intervention site and a standard based on the Harvard Standard reported in the literature (Nelson, 1969). Because of its current popularity, the National Center for International Health Statistics Standard supplied by the Center for Disease Control was used in some of the data sets as well (Hamill, 1976).

Second, in addition to the debate on local versus absolute standards,

there is the debate over sex-differentiated standards and unisex standards. Here, the difference is more than a parallel shift in standard. Particularly in cultures where male children are favored, use of a local sex-differentiated standard will mask the sex discrimination in the society by building it into the construction of the standard. Table 3-2 illustrates this phenomenon with the Candelaria-revisited data. The relationship of sex and nutritional status is determined using a Chi-square test for both the unisex Candelaria standard and the sex-differentiated Colombia standard. (The same classification was used in each.) As anticipated, the sex-differentiated standard shows no relationship; the unisex standard shows as significant relationship.

Third, there is no agreement with respect to the appropriate classification for identifying the malnourished. Although analysis can be done using the ratio of an observed value to a reference standard, a classification based on the ratio is far easier to use in the field (or in a clinic) and it provides a convenient method for presenting analytic results. Furthermore, a classification enables the intervenor and/or analyst to focus on those most at risk. For example, the Gomez weight/weight(age) standard is defined by a normal line based on observed weights for a presumably healthy population in Mexico and a four-way classification of nutritional status. Gomez called everyone whose weight/weight(age) ratio was eighty-six percent or more of the standard, normal; children falliing between seventy-six and eighty-five percent, grade I; children falling between sixty-one and seventy-five percent, grade II; and finally, children sixty percent and below, grade III (Gomez et al., 1955).

Gomez felt that these grades conveyed the sense of "at-riskness" in a useful summary form. The convenience of using a classification in the field arises because it is unnecessary to calculate the indicated ratio. It is sufficient to plot a child's weight on a growth chart that is already partitioned into the categories. Other approaches to classification include such things as the designation of children as malnourished if they are below one standard deviation of the norm.

Any analysis based on a categorization is highly dependent on the

TABLE 3-2

COMPARISON OF SEX-DIFFERENTIATED AND UNISEX STANDARDS - CANDELARIA REVISITED

Colombia Sex-Differentiated Standard

Sex \ Nutritional Status	Grades 2 and 3	Grade 1	Normal
Male	34 (32)	104 (102)	284 (288)
Female	30 (32)	101 (103)	297 (293)

Chi-Square = .54245: not significant at the .05 level

Candelaria

Sex \ Nutritional Status	Grades 2 and 3	Grade 1	Normal
Male	21 (27)	69 (84)	330 (309)
Female	34 (28)	100 (85)	293 (314)

Chi-Square = 10.899: significant at any level

categorization used. To illustrate how much the choice of both the "normal line" and "cut-points" can change just the simple rate of malnourishment, Table 3-3 is presented with four different combinations. The data are taken from the Esperanca intervention, which used the classical Gomez normal line and classification. We applied our version of the Harvard standard, which is more stringent, and the cut-points of 90%, 80%, and 70%. The percent of the population categorized as normal ranges from 33.2% to 63.6% while the percent labeled grade III ranges from .7% to 8.7%. (We should point out that one cause for the proliferation of standards is the desire of local areas to portray their malnutrition problem in such a manner as to justify increased funding. A mere shift in cut points or normal line can drastically change the severity of the malnourishment problem.)

A single dramatic illustration of the importance of the selection of the Standard and Classification can be drawn from the KOTTAR data. A local, unisex standard was taken from the growth chart used in the field for on-site diagnosis of problem cases. The cut-points of the classification were estimated from the diagram. Using this standard, the percent of children in the community suffering from second or third degree malnourishment dropped from 50.2% in January of 1975 to 42.5% in July of 1977. Using the NCHS/CDC standard and the Gomez classification, the malnourishment rate dropped from 50.4% to 45.5%--a difference of 4.9 points, as compared to the 7.7 points of the local standard. Note that the local standard was probably printed on the growth chart at a slight angle, causing older children to appear healthier than they really were. In the latter time period, there were more older children; therfore, the aforementioned "discrepancy" occurred.

We cannot resolve this quandary. More research to define "normal" must be undertaken. We have computed all our classifications using the locally applied cut-points and using the ninety, eighty, seventy percent with the Harvard Standard. But in order to simplify comparisons among different projects already analyzed by others, we have placed each of some fifty different standards in a data base management system which allows

TABLE 3-3

NUTRITIONAL STATUS IN ESPERANCA --
DIFFERENT NORMAL LINES AND CLASSIFICATIONS

NORMAL LINE		Grade 1	Grade 2	Grade 3	Normal	Total
	Classification(%)	30-60	61-75	76-85	86+	
GOMEZ	Number	6	85	210	526	827
	Percent of Total	.7	10.3	25.4	63.6	
	Classification(%)	30-60	61-75	76-85	86+	
HARVARD	Number	10	110	276	372	828
	Percent of Total	1.2	20.5	33.3	44.9	
	Classification(%)	30-69	70-79	80-89	90+	
GOMEZ	Number	34	134	233	426	827
	Percent of Total	4.1	16.2	28.2	51.5	
	Classification(%)	30-69	70-79	80-89	90+	
HARVARD	Number	72	224	257	275	828
	Percent of Total	8.7	27.1	31.0	33.2	

for the shifting of results from one standard to another.

Anthropometry Leads to Misclassification

An anthropometric measure, by its very definition, will misclassify some segment of the population. Consider the case of a weight/weight(age) standard. Such a standard is derived by measuring "normal" children of all ages to be covered by the standard. Usually, the median weight for all the "normal" children of a given age is selected as the reference standard. A curve is formed, using any of a variety of curve-fitting techniques, so that the standard appears to be continuous. The distribution of the observed weights about the curve for a given age is usually normal or bell-shaped. (In developed countries, there is often a skewness in the distribution; children in the uppermost percentile are often further from the median than children in the lower-most percentiles.) Some of these "normal" children will fall in the extremes on tails of the distribution and would be classified as malnourished if they were encountered in a developing country. In fact, in any study, some "normal" children will be identified as malnourished because they naturally fall into the extremes of the distribution. There is a built-in bias in the weight/weight(age) measure to classify small but healthy children as malnourished. Similar arguments can be made for the other anthropometric standards as well.

Conversely, some malnourished children in developing countries are no doubt misclassified because they are sufficiently big to appear normal relative to an anthropometric standard, while actually they are malnourished. Because there is no absolute measure of malnourishment to serve as a standard, it is impossible to estimate accurately the degree of error built into the use of anthropometric standards.

Also, child-age in months is a required datum for most anthropometric measurement, because nutritional status is computed by comparing an observed measurement with the reference standard for a given age. In many cultures there are problems associated with obtaining correct birthdates. This is further complicated by errors that can occur when calculating age in months. Errors in birthdate and/or age in month

calculations can occur for several reasons:

1. Lack of documentation of birthdate. Most children are born at home and births go unregistered. In some countries one has to pay to register one's child, and the cost is higher the older the child. Thus, there are situations where even the document may understate the child's age.

2. Incentives exist for misreporting a child's age so that he/she may remain eligible for program beneifts.

3. Program staff, health workers, etc., have problems converting age in years to age in months.

4. The protocol regarding the rounding up or down of months, depending on the number of days, is often unclear.

Anthropometric Measures May Not Respond to the Intervention

Although deficiencies with regard to height and/or weight in children are indicative of poor nutrition, an intervention which "works" may not result in an equal and opposite response in those measurements in all malnourished children.

First, one should expect that only some subset of a community is in a condition conducive to responding to an intervention.

> Indeed, in those rare studies where one has looked for a dose-response on performance, health and survival, through inproved nutrition in man, one finds a significantly lessened benefit as one improves nutritional state even at levels of the indicators universally accepted as inadequate. This means that for many indicators of performance, health and survival, one may not expect much improvement after intervention, unless the levels of the indicators in the malnourished population are quite different from normal levels in well-nourished regions. (Habicht and Butz, 1977; 138)

In the communities included in this study, many children were at or near normal at the outset. Because these children may not show any responses to an intervention, any aggregation over all children to show community responsiveness may mask the improvement in these children who were more deprived and more likely to respond.

The "intuitive" reaction to this phenomenon is to include only seriously malnourished children in the data set used for evaluation. However, for a variety or reasons to be expanded upon in the next chapter, there is a natural "regression toward the mean" in most collections of children. Concentration in analysis on only the most malnourished would require the separation of program impact from the "natural" recovery found in that subset of the population.

Second, there is evidence that some of the benefit of an intervention, particularly one that directly improves consumption or alleviates a situation causing existing consumption to be inadequate, manifests itself in ways other than greater growth.

> In the connotation of energy balance, the most logical explantion of the apparent inefficiency of the growth response is that a substantial part of the additional energy intake is applied to (a) a "deadaptation" process such as restoration of basal metabolic rates toward normal, and/or (b) an increase in physical activity probably expressed as play. (Beaton and Ghassemi, 1979; 35)

For any particular intervention strategy, it is not known at what point deadaptation ends and growth begins; therefore, one cannot state that that strategy is not working entirely on the basis of anthropometric measurements of growth.

Finally, stunted children may never exhibit the catch-up growth required to appear normal.

> The inference from the statistical findings in this study is that the chronic malnutrition, particularly during periods of rapid growth, results in irreversible growth stunting . . . If conditions improve, the child will be unable to catch up completely for all lost growth because biological time has passed and, with it, the opportunity for some growth. The catch-up that can take place is a function of the extent to which maturation was delayed and more "time" exists for growing. (Martorell, R., 1979; 387)

This is sufficiently important to merit detailed elaboration. Consider an individual for whom only weight and age have been collected for several points in time. Assume that the first of these measurements was made

when the child was eighteen montths old and that the weight/weight(age) of the child at this time was seventy percent. Assume that child remains at near seventy percnet of the standard throughout. The first response to this pattern is that the child was unaffected by the intervention. In fact, this conclusion **cannot** be drawn. Weight/weight(age) is a composite measure of chronic and acute malnourishment. If the child had suffered chronic malnourishment early in life (which may be nothing more than the aggregate effect of several bouts of acute malnourishment), he/she may be unable to regain the growth in skeletal structure (height) necessary to permit him/her to gain the weight that would improve his/her weight/weight(age) score. Thus, no change in percent of standard does not necessarily imply no program effect. An increase may well indicate an effect; a decrease does **not** indicate no effect.

The net result of this phenomenon is that for an individual, it is impossible to test for program effect using a limited time series of weight/weight(age) measures. A partial exception of this is the case where the time series begins at the time of the child's birth. A plot of the child's weight against age over time would then produce a complete trajectory for the child, indicating periods of acute malnourishment. From this chart, it may be possible to set a reasonable expectation for growth change, based on past history. For example, if a child's trajectory took a sharp dip just prior to the intervention (Figure 3-1), no change in percentile would mean no effect—prior stunting would not account for the lack of gain.

However, if there were no such dip, one could attribute the low weight/weight(age) score to either stunting or genetic difference. In this case, (Figure 3-2), we would expect no recuperative growth. Unfortunately, even in this case we cannot distinguish between stunting and genetically induced smallness; therefore, weight trajectory from birth is still only a partial solution.

Assuming that weights have not been routinely recorded prior to an intervention, the best way to assess the progress of a single child is through measurements of weight, age, and height. A low

FIGURE 3-1

INDIVIDUAL WEIGHT TRAJECTORY--ONE BOUT OF ACUTE MALNOURISHMENT

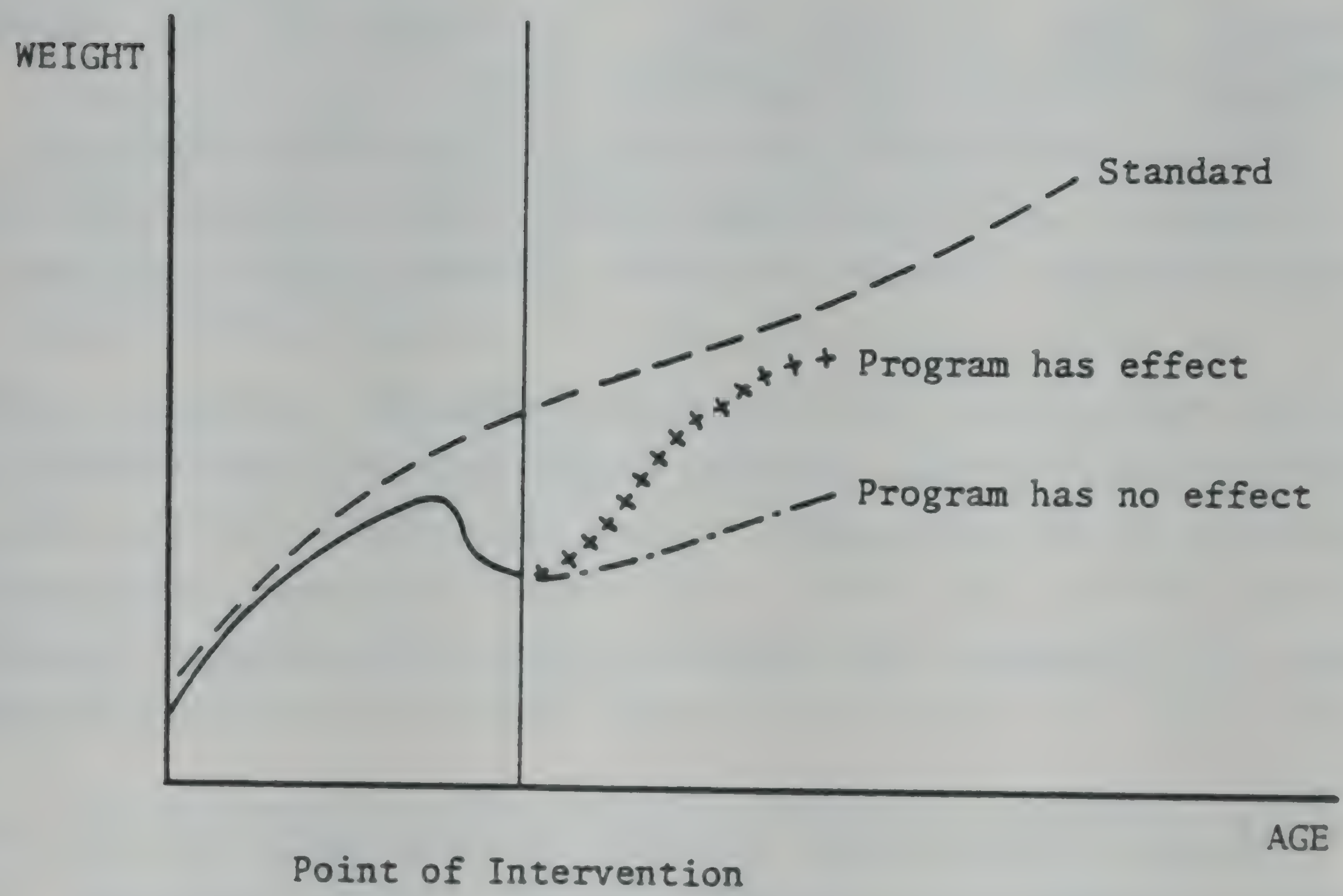
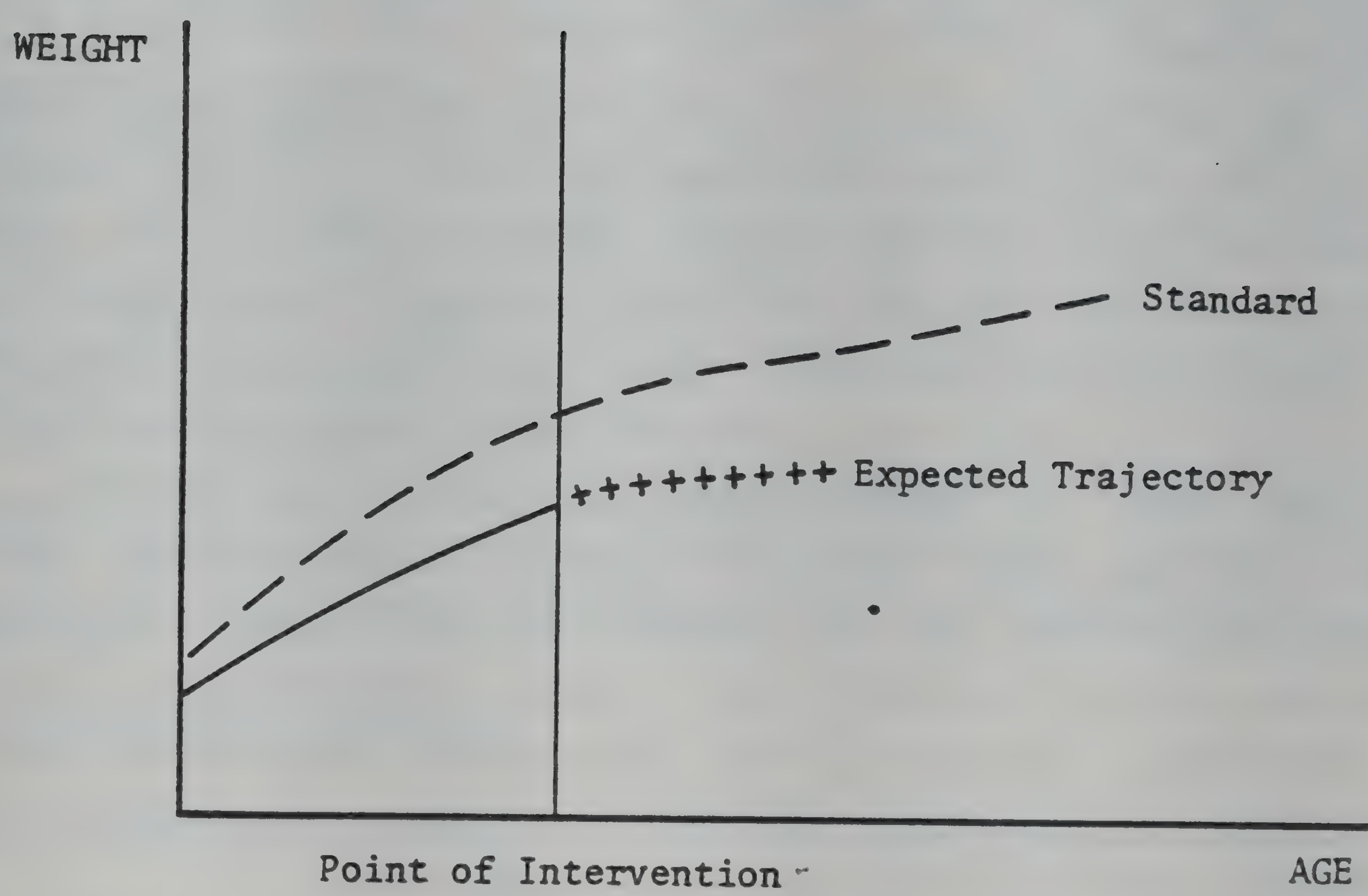


FIGURE 3-2

INDIVIDUAL WEIGHT TRAJECTORY--CHRONIC MALNOURISHMENT



weight/weight(age) reading coupled with a low height/height(age) score implies stunting. Furthermore, the weight/weight(height) ratio tells whether the child is currently in a wave of acute malnourishment in addition to this chronic condition. Again, however, if the child was stunted but scores well on weight/weight(height), we can **not** expect the intervention to produce growth. Thus, for an individual who has suffered chronic malnourishment, it is impossible to state that an intervention has failed. Of course, by recording height as well as weight, we cut down on the number of individuals for whom we do not expect growth; therefore, we recommend the use of height, as well as weight and age when field conditions allow.

In most of the interventions studied during this project, weight, height, and age were **not** gathered on children from birth. We must, therefore, consider the effect that this "missing" element will have on our analysis. The arguments have pertained to a **single** individual. What happens when we aggregate over all individuals? Unless every malnourished child at the start of the intervention had a history of chronic malnourishment and was currently normal in weight/weight(height), some children would improve. Thus, in the aggregate, we should see some decrease in the percent malnourished; however, this will understate the true effect of the program, since some children who benefit will not show a change in growth pattern.

In summary, the flaws in the anthropometric measurements are not trivial and, more importantly, they lead more often to an underestimation of change in nutritional status than to an exaggeration.

CHAPTER IV
FINDINGS CONCERNING DATA HANDLING

A major aspect of the effort to analyze community-level nutrition programs has been the acquisition, cleaning and analysis of longitudinal data sets generated as a part of those programs. Although experienced research social scientists are familiar with the procedures for developing and maintaining data, action-oriented intervenors in the third world are usually not proficient in this area. Furthermore, adequate technology and labor skill levels are generally not available in the field. Therefore, extra care must be taken in handling data in community settings where paper and pencils are uncommon and computers are science-fiction. In assessing our confidence in attribution for the separate case studies, the quality of the data played an important part.

We begin by enumerating several principles of design for collecting longitudinal data during the course of an intervention. These principles are a response to the problems encountered in our own analysis effort.

PRINCIPLES OF DESIGN

I. Know why you are collecting data <u>before</u> you begin.

There are many reasons for embarking on a data collection activity as part of a nutrition intervention and, in fact, many such "purposes" can be satisfied by a single data collection effort. However, the design of the data collection procedure must reflect the multiple purposes for its existence in order to achieve its goals.

We have seen data collected for six purposes:

1) Diagnosis - the identification of a target population (or individual) and the problems facing that population (or individual).

2) Education - the teaching of basic concepts of health and nutrition to service recipients through data gathering; for example, the education of mothers through the use of weight charts.

3) Management - the rationalization of decision-making by program leaders, particularly with respect to the allocation of resources to subsets of the target population, or between types of services made available to program beneficiaries.

4) Research - the enhancement of knowledge accruing to the "experts" with respect to the causes and cures of malnutrition.

5) External Evaluation - the determination by program backers (outsiders such as US/AID) of the relative merits of supporting or abandoning a particular program.

6) Bureaucratic Requirement - the meeting of requirements set by organizations controlling the resources flowing into a program.

Consider the potential conflict posed by bureaucratic requirement, external evaluation, education, and diagnosis with regard to the time interval for measuring the growth rate of children. Suppose a home-based supplementary feeding program (under PL 480) supplies a ration based on the number of malnourished children observed in the previous month. Thus, children must be weighed monthly to generate the required roster of beneficiaries. An external evaluation of such a program might require annual or semi-annual measurements of nutritional status. Education of the mother with regard to the importance of "putting on kilos" might call for quarterly visits—enough time for the mother to see a positive weight change. The diagnosis of at-risk children might call for monthly or quarterly measurements, depending on how often it is felt children must be checked to prevent catastrophic rapid decline.

This hypothetical situation did, in fact, exist in Honduras, where food supplementation was incorporated into a general Save the Children rural

development effort in the Pespire region. The bureaucratic requirement of generating lists of malnourished children to set the ration led to monthly measurements. The logistics of getting the children to the scale—a task that fell on the mother—were prohibitive. The program was later modified in response to the logistics problem and the realization that quarterly measurements were sufficient for diagnostic purposes. This modification coincided with a general overall review of the data gathering activity, a review that, according to Principle I, would have best preceded the first pass at data collection. Ultimately, the weighing program (and food supplementation program) were stopped, partly because the staff and the community were unable to overcome the image of chaos that formed during the early months of weighing.

Corollary Ia. **Make sure your field workers understand why you are collecting data.**

This corollary to Principle I is, perhaps, more important than the principle itself. In any event, it is more often overlooked in practice. In one of the interventions we studied, a baseline study of one village was executed with the help of outside consultants (one, a CSF staff member). An indigenous staff member of the project team was so impressed with the benefits of determining nutritional status of children through weighing that he initiated a study-wide weighing program; except that he forgot to record the weights. No one had explained the importance of observing changes in weights over time as a more sophisticated way to identify children who were at risk; therefore, the oversight.

The quality of the data--its precision, its completeness, and its timeliness—is a function of the commitment of the field staff to the data gathering activity. This commitment becomes more than a perfunctory doing of one's job only when the staff understands why they are collecting the data and, especially, when the staff **uses** the data to help the program beneficiaries.

II. Design the system for flexibility.

In Chapter VI, we develop the argument that one element contributing to the success of an intervention is a commitment to self-monitoring to guide program revision as the intervention matures and the local environment changes. The analysis of a longitudinal data base is one way to achieve such "informed incrementalism." Our experience suggests that attempts to review some portion of a longitudinal data set while an intervention matures invariably suggest additional changes for the data collection protocol. Also, changes in the program itself often suggest modifications of the data base. A good data system will be able to respond to such suggestions.

In an intervention studied in which children were observed periodically, an experiment using arm bracelets on children as a surrogate for weight measures of nutritional status was undertaken. However, forms for recording collected data as well as the formats for making the data machine readable were unable to accommodate the experiment conveniently. The outcome was a haphazard attempt to append another column (or two) onto each card with the results of the "bracelet" experiment. The ensuing data proved unanalyzable; because of the inconsistent and incomplete use of one or two additional columns, it was never known who actually participated in the experiment.

An ongoing data collection process will assuredly require the recording of special variables or new variables as experiments are tried, and/or environmental conditions change. The system, whether machine oriented or manual, should easily adapt to such changes. Data forms and formats need not be etched in rock; they, too, must grow during an intervention.

III. Feed back results of data gathering to field personnel and management.

We made reference earlier to the positive role feedback plays in contributing to the commitment to the staff. This is sufficiently important to become a principle in itself.

In one intervention, the number of simple recording and/or keypunch

errors resulting in a data value being out of range increased from 4.9% to 15.7%, to 16%, to 18.4% with each distinct batch of collected data. Yet feedback designed to restore accuracy was not given to either the collectors of the data or the data processing people.

Feedback not only contributes to commitment; it shows where errors are made. Thus, a table ordering program participants by weight as a percentage of weight/weight(age) standard shows immediately some "unbelievable" values. For the Indonesia intervention, such a table was generated to check on the stated ages of children whose calculated percent of standard were unusually high or low.

Finally, feedback boosts the morale of field workers by demonstrating results. In several sets, our feedback triggered positive response in field workers simply because it meant that someone "really cared." The results were less important than the notion that the data were being used.

PRINCIPLES OF IMPLEMENTATION

I. Collect unambiguous data.

This project received most of its data after it had been collected and, in most cases, there were sources of ambiguity. In some cases, the ambiguity could be resolved through hard work; in others, the ambiguous data were useless. The causes of ambiguity are many: a poorly designed questionnaire, an awkward recording form, a badly conceived coding scheme. Even a well-executed data gathering scheme goes awry if field personnel are haphazardly trained in its application.

In one data set, weight measurements were taken in pounds at the start of the program but, somewhere in the middle of round 2, the units of measurement changed to kilograms. (In this case, we were able to convert all the earlier measurements to conform to the others by testing their logical consistency with the time series for each child, and making a judgment as to the correct unit.) In another instance, food intake measurements were recorded by some interviewers in grams while other interviewers used kilograms. (Again, judgments were painstakingly made to

make the data uniform.)

An example of a data collection form that stymied interpretation is one that asked for the location of consultations made for curative health services by visit: first visit, second visit, third visit, etc. Our first assumption was that there could be no second visit during a round without there being a first during that same round. However, there were many cases where the data indicated the reverse. The errors may have been in our assumption, in inconsistent interpretation by field workers, in coding, in keypunching, or in all of these. We tested other assumptions as well: for example, that a second visit could not occur without there being a first during the time of the program. None explained the data, and it remains virtually useless.

Our examples thus far seem to illustrate nothing more than sloppy work; however, there are some sources of ambiguity that even the most cautious efforts may not preclude. In particular, the ambiguity problem proves to be exceptionally thorny with regard to identification of the program participants and their relationships. In computerized systems, an identification number is usually created for program participants. Typically this number tries to capture the family, the household, and even the village of the participating child. Ambiguity creeps in during efforts to define these terms.

Consider the simple concept of birth order. It has two immediate interpretations: order of the child among living children in a family, and order of a child among all children given birth by a single mother. Both concepts have potential theoretical importance. Malnourishment may be a function of the number of mouths that must be fed (the first measure) as well as the health status and pregnancy history of the mother (the second measure). Even if both measures are made, what of twins, adoptions, cousins living with the family, children of former marriages? In multiple-family households, the relevant concept may be child's birth order with respect to all children in the household.

When analyzing data relating households, families, individuals, and nutritional status of children, the definitions of these concepts are of

critical importance yet are extremely difficult to derive. The level of detail needed to capture all the instances is excessive and, even if the definitions are precise in the "sterile" office of the systems analyst, they are quite difficult to implement in a questionnaire in the field. We have an instance where a child's birth order was asked in each round. Many children experienced widely fluctuating birth orders over the course of the study. In another instance, every member of the household was uniquely identified, yet the problem of matching which child belongs to which mother in multiple-mother households was overlooked.

Good definitions and clear questions are the keys to avoiding ambiguity. Field testing the vehicle (questionnaire, form, etc.) used to collect data is the best insurance of clarity. Also, advance planning with respect to analysis (rather than retrospective data manipulation) identifies potential sources of ambiguity and/or misinterpretation.

II. Do not eschew redundancy in identifying program participants.

All data collection methodologies, manual or computerized, use some form of identification for program participants. In manual systems, human judgment often uncovers cases of mistaken identity. In computerized systems, such fortuitous discoveries are not made until the computer stumbles upon impossible situations. In interventions where repeated observations are made on individuals over a substantial period of time, mistakes in identification result in children who grow younger, change sex, and experience weight changes that defy all theories of human growth. The best way to avoid such errors--or at least provide a means of correcting them--is to identify everyone twice. For example, in Pespire, SAVE incorporated the child's sex and birth date on each observation form even though these data elements were invariant. Mistakes in ID numbers were corrected by matching sex and birth date as well.

Unfortunately, we also had examples where a desire for "efficiency" eliminated all such redundancy. In one case the original data collection form had sex and birth date but, to save writing and keypunching, these

were eliminated in a modified form. The end result was hours and hours of rectifying impossible matches with, in many cases, the elimination of "unmatched" observations.

III. Record basic data elements, not just derived variables.

It is wisest to measure and record basic data elements in the field and categorize, transform, or otherwise derive other variables in a more favorable setting (or a computer). To illustrate this principle, consider the determination of a child's age. If a child's birth date is known, age in months can be calculated; however, the calculation is not trivial. Even if the field person collecting the data needs to know age (to determine nutritional status as a step in prescribing treatment), it is wisest to record birth date and observation date and recompute age in the office or in a computer. During one of our field visits, it was discovered that approximately thirty percent of all ages were being calculated incorrectly by field staff. The team of "experts" conducting the field visit recalculated all ages in the field and, after recomputation in a more serene environment, it was discovered that they made errors in fifteen percent of the calculations. It may seem to be "make work" or a form of redundancy to record the basic data when the derived variable is of primary interest, but the time and expense saved by not recording basic data elements is more than lost in subsequent error correction and data loss.

In the nutrition field, the same can be said for such derived variables as nutritional status, percent of standard, change in weight between visits, etc. Although it might be desirable to compute each of these variables in the field for use in on-the-spot diagnosis, the raw data used in computing each of these variables should be recorded to allow for corrections in a more convenient setting at some later date.

Another form of "data loss" is experienced (especially in surveys) when a continuous variable is categorized prior to its being recorded. Consider, for example, the age of a child's mother. For many purposes, it would be sufficient to know in what five-year category a particular mother falls.

Yet, it is best to record actual age--an element which can then be recategorized or used as is in analytic procedures requiring continuous variables. As a case in point, a baseline survey we looked at was done almost entirely in categorized variables to facilitate hand tabulations of the results. Later, when we put the data into a computer, the losses resulting from the categorizations were felt in questions such as "How many total people (adults and children) live in this house at the present time: zero to three, four to six, seven to nine, ten to twelve, thirteen to fifteen, sixteen to eighteen, or greater than eighteen?" Here, a family with four and a family with nine are only one category apart and are "ranked" the same as two families, one with six people and the other with seven. This loss of discrimination is easily avoided by recording the number of people directly.

IV. **Record something for every possible entry.**

Two extremely frustrating problems encountered during our efforts to clean data have been: (1) distinguishing a response of "0" or "none" from missing data; and (2) correcting keypunch errors resulting from keying data in the wrong columns. Both of these problems can be minimized by using a unique code for both missing data and "not-applicable" data.

In one nutrition survey, time series observations for individual children were recorded by entering the month of the observation (2 columns on a card), the year of the observation (1 column), the child's weight (3 columns), the health status of the child (1 column), and two entries for type of illness if sick (1 column each). If a child was healthy, both the first and second "illness type" fields were left blank. Hundreds of interviews had at least one instance in which the keypuncher forgot to leave the blank spaces. Also, many observations indicated a health status of "sick" but made no entry in the "illness type" fields. In the first case (omission of blanks), subsequent observations were all "wrong." In the latter case, it cannot be determined whether the child was really healthy (an error in "health status") or whether the child was sick (omission of "illness type"). Both problems would have occurred much less frequently

had a not-applicable entry been made for "illness type" when the child was healthy.

V. Check for and correct errors as close—in time and space—to the source as possible.

The sooner errors are found, the easier they are to correct. Ideally, errors would be detected on-site within days if not hours of the time the data are collected. This would facilitate correction (it might even be possible to retake suspect measurements) and keep errors from proliferating.

Recognizing that data collection in remote areas might preclude on-site checking, it is still best to find and, if possible, correct errors as close to the source of data as possible. With the exception of two of the data sets, none were checked at all until we received the data. The Thailand data set was cleaned by a fairly complex procedure of error checking in Boston, correcting back in Thailand, and joint correction of computer cards in both locations. The data set from Esperanca had some field checking (our staff was involved in the on-site coding of the baseline study and supervised the entire revisitation survey), and errors found in Ann Arbor were made known in the field for correction on-site.

The remaining data appear to have been unchecked, and the expense of our correcting the data so far removed from the field was outrageously high. The amount of data lost because we were unable to make a correction was also considerable. What troubles us most is that much of this "obviously" incorrect data was used in the field for diagnostic purposes. Of course, the errors that we find may be coding and/or keypunch errors; the field measurement may have been accurate. Yet, it is distressing to find a high percentage of time-series errors where each individual observation could be correct, but where the sequence is impossible. In one data set, we had countless incidents where the recorded age of the child did not change by the amount of time between observations. Clearly, consistency checks should be made at the time of the observation to guarantee that prescribed treatment is appropriate.

VI. Verify all keying in a machine-based system.

In computer-based data management systems, project leaders apparently underestimate the cost of getting data into machine-readable form. We presume that these costs are underestimated because of the number of projects that "save money by not verifying their data entry work." (Verification is the process by which keypunchers are checked by other keypunchers to find pure keypunch errors.) In one of our larger data sets, a third of the cards had some sort of keying error—errors that could have been caught more efficiently and at less expense by a keypunch operator than by a highly paid systems analyst.

As a postscript, we return to Design Principle III—feed back results to field personnel and management. The single most upsetting characteristic shared by almost all of the data that played a part in this research was its lack of productive use in governing an intervention. This "disuse" leads to the violation of all of our implementation principles: ambiguities are never found, the identification of participants is unnecessary, derived variables are never found to be in error, missing entries are not at all troublesome, errors are never corrected, and verification is irrelevant. Use of the data, by contrast, makes these principles obvious--not just to the analyst but to the change agent in the field.

PROCESS OF DATA CLEANING

There are essentially three types of errors to be corrected in cleaning data: (1) entries that are out of legal range; (2) inconsistent entries within a single observation; and (3) entries that, although they appear to be correct within an observation, are inconsistent with entries that have preceded it.

Examples of the first type of error are a 12-month-old child who weighs 80 kilograms (a "normal" 12-month-old weighs on the order of 8 kilograms) or a sex reported as "3" when legal codes are "1" for male and "2" for female. Examples of the second type of error are instances where a health status indicator might be "well" at the same time a disease type indicator says "measles," or a mother says **No** when asked if **any** children

have died, yet when asked **how many** have died answers **two.** Examples of the third type of error are a child whose birth date post-dates an observation date or whose weight ranges from 7 kilos to 7.2 kilos to 4.3 kilos to 7.5 kilos over short time intervals.

Each type of error can be generated by measurement problems, coding or transcription problems, and/or keypunch problems. Unfortunately, unless the original data collection forms are available, there is usually no way to know how an error came about. (Some keypunch problems are "obvious," such as an entire card shifted one column.) Whatever the cause, analysis requires the elimination or correction of spurious values.

In general, the first two types of error—range violation or inconsistency within observation--must be located and eliminated prior to locating consistency errors between rounds in a time series. This is true because many range violations and within-observation inconsistencies appear as between-round inconsistencies. The correction of error type one or two automatically corrects the error type three. Since the latter are most difficult to correct, it behooves the analyst to begin corrections with error categories one and two.

Consequently, the first step in error correction is the identification of range violations and within-observation inconsistencies. Because each of our data sets had a unique format, unique programs had to be developed to perform all such error searches. Fortunately, many of the "subroutines" used to check particular variables were transferable between unique master (mainline) programs. For example, routines to compute time in months between given dates (birth date and observation date), and to determine whether a child of a given weight and age is in a reasonable percentile relative to a standard, were used in all of our data sets.

Once located, range check violations and within-observation inconsistencies must be corrected or eliminated. Elimination is the simplest course of action. In fact, it is possible to set errors to a "missing code" automatically in the range check programs themselves. In most cases, we shunned this approach. It was felt that errors probably appear in atypical situations more frequently than in run-of-the-mill

instances. Thus, elimination of errors is tantamount to biasing the data set to the norm by eliminating extremes. We opted for a strategy of examining all errors, and never eliminated an observation or a variable within an observation until human judgment determined it to be the only recourse. Most errors can be corrected, even without referring back to original data sources.

The following examples illustrate how this was done. The corrected values for some errors are obvious. In one data set where weight was measured in grams, a series of observations were in the 70,000 to 99,000 gram range—apparently a consistent mispunch or miscode of weights in the 7,000 to 9,900 range. Still other range violations are less obvious and require more subjective judgment. Correction in these cases is optional. In another data set, "0" was used to indicate a healthy child, "1" indicated sick, and "2" meant hospitalized. No other values were legitimate; yet, there were quite a few "3" entries. This particular data set was keypunched and not verified. The "3" on a keypunch is the same key as the letter "O." The number "0" is right above the letter "O." A keypuncher holding the numeric shift key while mistakenly hitting the letter "O" instead of the number "0" will generate the number "3." The multiple occurrences of the illegal "3" code in the absence of other erroneous numeric entries strongly suggest this keypunch error. Thus, the "3" entries can be reasonably converted to "0."

Consistency errors within an observation can occasionally be corrected in similar fashion. In one of our data sets, a yes/no code was used for each ossification center in a hand x-ray. Also, the total was recorded. Some of the time, however, the total did not agree with the sum of the "yes" responses recorded for individual centers. These could often be made to agree by changing the recorded total. Similarly, in another data set, a question regarding land ownership called for a "1" if the respondent owned land and "2" if the respondent rented. This question was followed by an inquiry into amounts of each, subdivided into wet and dry. Often, respondents claimed ownership while claiming rented acreage of one type or another and vice versa. In most of these instances, the respondent

really held **both** types of land. A new code was created—"3"—for both ownership and rental.

Once a set of time series data is known to consist of "clean" individual observations, consistency between observations for a single child can be checked. Here, by examining the time series, it is often possible to make judgmental decisions concerning corrections. A child who was male in 5 of 6 observations was probably male for the sixth. Similarly, weight sequences like 7.2, 7.5, 4.7, 7.8 kilograms can safely be assumed to be 7.2, 7.5, 7.7, 7.8.

Most troublesome among consistency problems between rounds are errors in identification numbers; that is, the inconsistency arises because the time series is not really for the same child. We have already discussed the need for redundancy of identification, yet the argument for designed redundancy must be repeated, because little can be done in data cleaning besides deleting the inconsistent entry. Even though it is known that the entry belongs to another child, there is no way to determine which one.

In almost all settings, there are three types of data that are relevant: constants, or data which do not change (a child's sex or birthdate); semiconstants, or data which usually stay the same but over a long period of time may change (a family's address or income); and variables, or data which change on a more or less continuous basis (a child's weight or health status). A good system encompasses the three data types.

Ideally, upon program initiation all three types will be collected. In essense, this is a baseline survey of the community at the start of a program, and it is essential that all newcomers (either by migration, births or program expansion) are surveyed at the time of their first participation in the program. The design of such a baseline survey is not a trivial matter. It is important to note one major complicating factor: some data are individual-specific, some are relevant at the family level, and other data are best viewed from the household's perspective. The "baseline survey" should ameliorate this complicating feature.

Baseline data provide the foundation upon which to establish a continuous monitoring system. If children are to be weighed and/or their

health monitored over time, a record of each observation should be made. Because this data is frequently collected, it should be kept to a minimum quantity to keep staff workloads at a reasonable level. Some constants from the baseline should be included to guarantee proper pairing of the incoming data to the existing data set.

Finally, a "resurvey" should be administered periodically (yearly) to learn of changes in the semiconstant data. A cleverly designed system might require only **changes** be added. For example, in a computerized system, the computer might generate the questionnaire with the previous answers indicated. Then only changes need be noted. In a manual system, the old form may be used to indicate changes over time.

To make this system work, rapid feedback of the data must be made available at all levels to: field persons, management, and the mothers and children themselves.

A "living" system such as this would need to have built in error checking devices. In manual systems, certain entries should be checked against previous entries at the time the data are permanently recorded, and prior to generating tallies for reporting purposes. In computerized systems, all entries should be checked against the computer data bank. We advocate that the historic data base should be assumed correct until a new and inconsistent entry is proven to supersede an earlier entry. Thus, a new observation is checked against the data base. If it is correct, it is added; if it is wrong, it is held out for verification.

The design and implementation of such a system may seem far-fetched to those familiar with the realities of interventions in developing countries. To run such a system "manually" suggests scaling down its size to make it manageable. To use computers in places like Aceh or Santarem may seem impossible, while shipping data to the United States and shipping reports back may be expensive and inefficient.

Yet the use of computers in Aceh or Santarem is really quite feasible. Computer technology is such that for a few thousand dollars, an extremely useful self-contained micro-processor can be obtained. Micro-processors are portable, reliable, and could be used as intelligent data entry terminals

in the field. If the intervention is not in too remote a setting or is not too big, this hardware could be used for primary storage and report generating. Importantly, information stored on this equipment is readily transferable; most major computer facilities have the capability of reading data stored on the cassettes or diskettes used by today's micro-processor. Thus, major "number crunching" and analysis could be done off-site on bigger machines if necessary.

In summary, one could "register" all participants in an intervention with a survey type device to get family and household data. Range checks and within-observation consistency checks would be made in the field. If possible (if the number of participants is not too great), consistency between rounds could be field checked also.

In larger time intervals, resurveys could be completed to learn of changes in socioeconomic status, family composition, etc. The local economy, weather conditions, and other environmental indicators could be monitored periodically as well.

Reports would be generated and the capability for producing "special studies" developed as the system grew. Not only would intervention managers monitor the program, but they would also assess the usefulness of the system and make modifications as required.

CHAPTER V
FINDINGS CONCERNING ANALYTIC METHODOLOGY

In the last two chapters, we have identified and discussed two issues which lessen the discriminatory power of analyses of nutrition programs using anthropometry. First, the most often used anthropometric measurements were shown to lack sensitivity to intervention in many common situations. Second, the data itself were identified as lacking in completeness and precision in most service-oriented interventions. Unfortunately, several issues associated with the analysis methodology itself further complicate most efforts to make conclusive recommendations from the data. In this chapter, we will identify three such issues which are largely ignored in the literature on nutrition evaluation. We will then discuss two analytic procedures we use to minimize the negative aspects of these issues, and finally we will summarize our overall process of analysis.

THE CONSISTENT UNDERSTATEMENT OF RESULTS

In Chapter II, we reported the results of eight in-depth case studies. The single most important computation in each case study led to a comparison of the percentage of malnourished children on site at two points in time or, when control populations existed, in matched groups receiving different levels of service. There were moderate to substantial reductions in the percentage of malnourished in each case and, in several cases, observed reductions were attributable to the intervention. Given the mechanical process by which such comparisons are made, there is a **consistent understatement** of the differences between groups. The understatement of differences is a result of the misclassification of some children as to nutritional status--the magnitude of the understatement is directly proportional to the magnitude of the misclassification.

Note, this understatement phenomenon is not limited to nutritional

studies alone. Any study relying on a comparison of subjects falling into classes will also encounter understatement of change. For example, a study to evaluate the effectiveness of an adult education program in erasing illiteracy where entering students are classified as literate or illiterate and retested upon completion of the program will understate results. However, in such a study, it can be assumed that the test used to classify the students is accurate—therefore, understatement would be nil. We argued in Chapter III that such an assumption cannot be made when anthropometrics are used to determine malnourishment. Normal children who are genetically small, and malnourished children who are large boned and heavy will be misclassified even when their weights, heights and ages are accurately recorded. Misclassification due to misstated ages, misread weights, and poor scales compounds the problem even further.

Consider a simple hypothetical example of a successful nutrition intervention. To keep the mechanics simple, assume that 200 children are participating in the intervention and that they are classified as normal or malnourished at both the outset and at the conclusion of the program. Traditionally, observers would state the percent malnourished at the outset, the percent at the conclusion, compare the percentages and draw a conclusion with regard to the efficacy of the intervention. Table 5-1 presents the numerical example. We assume that 15% of all children are misclassified. In reality, we cannot know the "TRUE" state—we only know the "OBSERVED" state. In this example, malnourishment dropped from 45% to 40% in the "TRUE" state, a reduction of 11.1%. We observe a drop from 46.5% to 43%, a reduction of 7.5%. The numbers in the "OBSERVED" half of the table are computed from the "TRUE" numbers under the assumption of 15% misclassification. For example, the 107 "OBSERVED" normal children at the outset is 85% of the "TRUE" normal plus 15% of the "TRUE" malnourished (110 x .85 + 90 x .15 = 107).

A word of warning for the reader is in order. In examining this example, the tendency is to believe that it is contrived to produce the startling result that impact was understated. One might believe that by altering the rates of misclassification or the rates of improvement in the

TABLE 5-1

ILLUSTRATION OF UNDERSTATEMENT--BEFORE AND AFTER DESIGN

	"TRUE" INITIAL STATE		"OBSERVED" INITIAL STATE	
NORMAL	110 (55%)	NORMAL	107 (53.5%)	(110 x .85 + 90 x .15 = 107)
MALNOURISHED	90 (45%)	MALNOURISHED	93 (46.5%)	(90 x .85 + 110 x .15 = 93)
	200 (100%)		200 (100%)	

	"TRUE" FINAL STATE		"OBSERVED" FINAL STATE	
NORMAL	120 (60%)	NORMAL	114 (57%)	(120 x .85 + 80 x .15 = 114)
MALNOURISHED	80 (40%)	MALNOURISHED	86 (43%)	(80 x .85 + 120 x .15 = 86)
	200 (100%)		200 (100%)	

TRUE PROGRAM EFFECT--Malnourishment Dropped From 45% to 40%, an 11.1% Reduction

OBSERVED PROGRAM EFFECT--Malnourishment Dropped From 46.5% to 43%, a 7.5% Reduction

"TRUE" case, the direction of the error can be reversed; that is, an overstatement could be demonstrated. This is not so. Even if we only misclassify some percentage of the malnourished as normal will we have understatement—and vice versa. Even if malnourishment gets worse, we will have an understatement. As long as we rely on a classification and have some built-in source of error in that classification, we will have understatement. Because we do not know the "TRUE" state in the real world, we can never know the magnitude of the misclassification.

Table 5-2 presents a somewhat more complicated example—a before and after research design using controls. To illustrate the increasing magnitude of understatement in response to poorer classification, we assume two different misclassification rates—fifteen percent and thirty percent. The summary figure used to define net effect in this design is the net change in the difference in malnourishment between participants and nonparticipants. Again, there is dramatic understatement of net effect—by misclassifying three out of ten children we understate the improvement due to intervention by more than half.

The numbers used in our hypothetical example above were not selected arbitrarily. There is good evidence to suggest that the minimum misclassification with respect to nutritional status when using weight/weight(age) is approximately fifteen percent. In the field, we can expect even greater misclassification—even as high as thirty percent.

The policy implications of this consistent understatement of results are many. Most obvious, a comparison of the efficacy of two or more programs with different but unknown degrees of understatement is extremely risky—outcome differences may be nothing more than differences in the accuracy of the classifications used. To guard against this, at the very least there must be a standardization of classification procedures to keep constant the degree of understatement among the alternatives under consideration. For instance, if measures for incidence or prevalence of diarrhea are less reliable than measures of respiratory infection, then a diarrhea control intervention will appear to have less effect than the respiratory infection, even if both are equally beneficial. In this example,

TABLE 5-2

ILLUSTRATION OF UNDERSTATEMENT--CONTROLS

TRUE INITIAL STATE

	Participants	Nonparticipants
NORMAL	110 (55%)	100 (50%)
MALNOURISHED	90 (45%)	100 (50%)
	200 (100%)	200 (100%)

TRUE FINAL STATE

	Participants	Nonparticipants
NORMAL	160 (80%)	80 (40%)
MALNOURISHED	40 (20%)	120 (60%)
	200 (100%)	200 (100%)

TRUE PROGRAM EFFECT--35% REDUCTION IN MALNOURISHMENT

MEASURED OUTCOME ASSUMING 15% MISCLASSIFICATION

	Participants	Nonparticipants
NORMAL	107 (53.5%)	100 (50%)
MALNOURISHED	93 (46.5%)	100 (50%)
	200 (100%)	200 (100%)

	Participants	Nonparticipants
NORMAL	142 (71%)	86 (43%)
MALNOURISHED	58 (29%)	114 (57%)
	200 (100%)	200 (100%)

OBSERVED PROGRAM EFFECT--24.5% REDUCTION IN MALNOURISHMENT

MEASURED OUTCOME ASSUMING 30% MISCLASSIFICATION

	Participants	Nonparticipants
NORMAL	104 (52%)	100 (50%)
MALNOURISHED	96 (48%)	100 (50%)
	200 (100%)	200 (100%)

	Participants	Nonparticipants
NORMAL	124 (62%)	92 (46%)
	76 (38%)	108 (54%)
	200 (100%)	200 (100%)

OBSERVED PROGRAM EFFECT--14% REDUCTION IN MALNOURISHMENT

* Program effect is the net change in the difference between participants and nonparticipants--before to after

the reported relative effectiveness is determined by the reliability of the measures used.

The same phenomenon is also true within a given measure when changes in reliability occur over time. If a nutrition project is monitoring nutritional status over time and for some reason a change in data gathering procedures decreases the reliability of the measures, it will appear that the nutrition project is less effective than it used to be, simply because the reliability decreased.

THE NECESSITY OF ACCOUNTING FOR AGE

By definition, the computation of change in nutritional status requires that data exist at two (or more) points in time. Although many variables may stay relatively constant over time, at least one variable, the ages of the children involved in the program, must change. If malnutrition were uniformly distributed over all ages, this would not be troublesome. However, in most cases such a uniform distribution does not exist. There is often a critical period in each child's life, some six months after weaning, when the combined effects of insufficient consumption and increased exposure to disease-causing agents (especially diarrhea) lead to a period of malnourishment. In the absence of intervention, some children will recover due to the natural development of antibodies against disease and the improved ability to consume solid foods. In the aggregate for a society, this phenomenon causes the non-uniform distribution of malnourishment. Typically, those age groups between eighteen months and three years (depending on culture) suffer most from the ravages of malnutrition. Again in the aggregate, death of the most severely malnourished can eliminate the bottom tail of the distribution accentuating the nonuniform nature of that distribution.

As a consequence of this nonuniform distribution of malnourishment in the preschool population, we have envisioned a "least cost" intervention. Measure a group of children ages one to four to determine their nutritional status. Measure those same children two years later. The natural recovery process will result in a smaller percentage of children being

malnourished--not because of any "real" improvement but because the children are all older.

Let us demonstrate this with an example drawn from the Save The Children program in Pespire, Honduras. Table 5-3 gives the counts and percentages of children who are normal and malnourished at the start of the program (malnourished children are limited to those falling into grade II or III as determined by weight/weight(age) using a version of the Harvard Standard employed by the SAVE staff in Pespire). There were 250 children, but weights were recorded on only 242 of them.

We observe immediately that 28.1% of the children were malnourished and that we do see a nonuniform distribution of malnourishment over age groups. (The aberration of the high percentage of children in the sixth year category is a result of a program selection bias to include only those older children who were still "in trouble.")

Table 5-4 presents a similar set of data for the fifth round of weighings which took place about 13 or 14 months later. This table is for the same 250 children—**not** for all participants in round five. Only 183 of the children who were still less than 72 months old were weighed; therefore, the counts are somewhat smaller.

The table shows that malnutrition declines--from 28.1% to 25.7%. However, a closer look reveals how misleading that "gross" decline can be. The one-year-olds at the start of the program exhibited a malnourishment rate of 16.3%. Most of those one-year-olds had turned 2, 13 months later (some of the oldest were 3), and their malnourishment rate had increased to 25.9%—as expected due to aging. Similarly, the 3-year-olds who were now 4 showed a marked decline from 40% to 14.6%—also as expected due to aging. The age of most risk in Honduras appears to be between 2 and 3. By the time of the second measurement, most of the children had passed this age of most risk; therefore, we see an overall decline in malnourishment.

The importance of age in determining the risk level of children is widely accepted. What is not accepted is the importance of accounting for age shifts in a population in doing "before and after" types of analysis.

TABLE 5-3

NUTRITIONAL STATUS OF THE POPULATION AT THE START OF THE PROGRAM, PESPIRE, HONDURAS

NUTRITIONAL STATUS	AGE						
	One	Two	Three	Four	Five	Six	Total
NORMAL	41 (83.7)	29 (72.5)	30 (60.0)	29 (61.7)	40 (83.3)	5 (62.5)	174 (71.9)
MALNOURISHED	8 (16.3)	11 (27.5)	20 (40.0)	18 (38.3)	8 (16.7)	3 (37.5)	68 (28.1)
TOTAL	49	40	50	47	48	8	242

TABLE 5-4

NUTRITIONAL STATUS OF THE SAME CHILDREN THIRTEEN MONTHS LATER, PESPIRE, HONDURAS

NUTRITIONAL STATUS	AGE						
	One	Two	Three	Four	Five	Six	Total
NORMAL		27 (71.1)	22 (73.3)	35 (28.4)	27 (69.2)	25 (71.4)	136 (74.3)
MALNOURISHED		11 (28.9)	8 (2617)	6 (14.6)	12 (30.8)	10 (28.6)	47 (25.7)
TOTAL		38	30	41	39	15	183

See, for example, the Synectics Corporation report, "Application of a Field Guide for Evaluation of Nutrition Education to Programs in the Philippines," pages 39-51, where pre and post measurements of a population participating in Mothercraft Nutritional Centers were used in an analysis that did not account for the natural aging process described above (Jones and Munger, 1978).

THE CHARACTERISTIC CURVE

The phenomenon of children passing through a stage of high risk after weaning is just about universal. The phenomenon is so important that we chose to construct a device for demonstrating it and for learning from it. We call the device a **characteristic curve**—a graphic representation of the relationship between malnourishment and age in a population. Figure 5-1 is the curve for the Pespire, Honduras, data in Table 5-3.

The shape of this particular curve is typical in most (but not all) environments with a prevalence of malnourishment. As expected, there is a low initial rate of malnourishment which climbs to some peak after weaning and drops off again as children reach the age of five. We became so used to seeing curves of this type that we were somewhat surprised to find curves that did not conform to this general model. Consider the curve for Pakistan derived from the CARE five-country study (Figure 5-2).

This curve approximates a straight line—there does not appear to be a recovery phase. The relatively high value for children between seven and twelve months is either an aberration or a very revealing data point. If the value is correct, it indicates a peak age of "at riskness" in the first six months of life, and the entire curve represents a period of very modest recovery. This situation could be found in areas where mothers are extremely malnourished and breastfeeding does not give sufficient nutrients to the child, or breastfeeding is stopped very early (for example, at 3 months). Food intake in later years would be insufficient to enable much recovery.

Our speculation of the causes of a particular shape of a characteristic

FIGURE 5-1

CHARACTERISTIC CURVE--PESPIRE, HONDURAS--ROUND 1

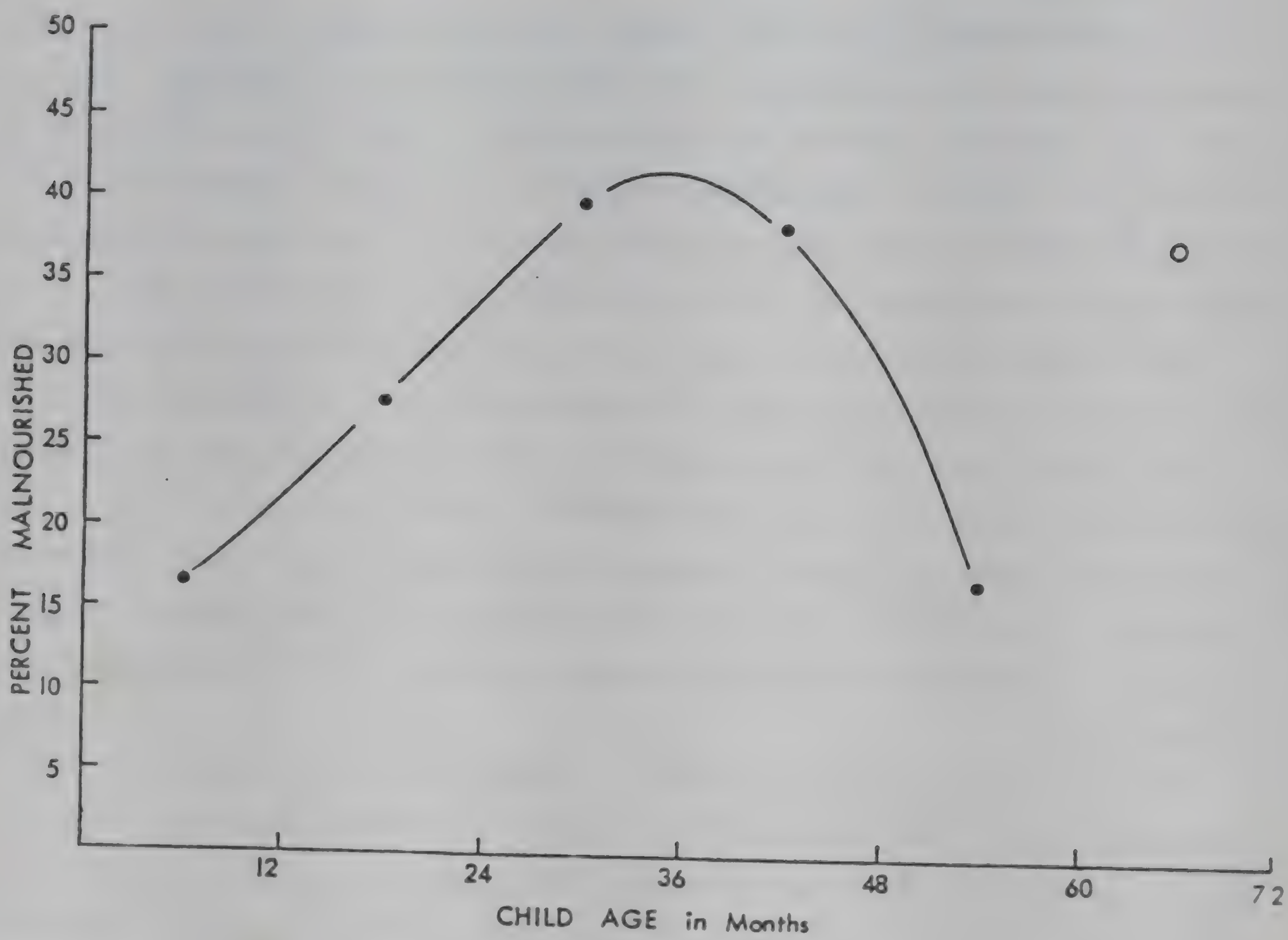
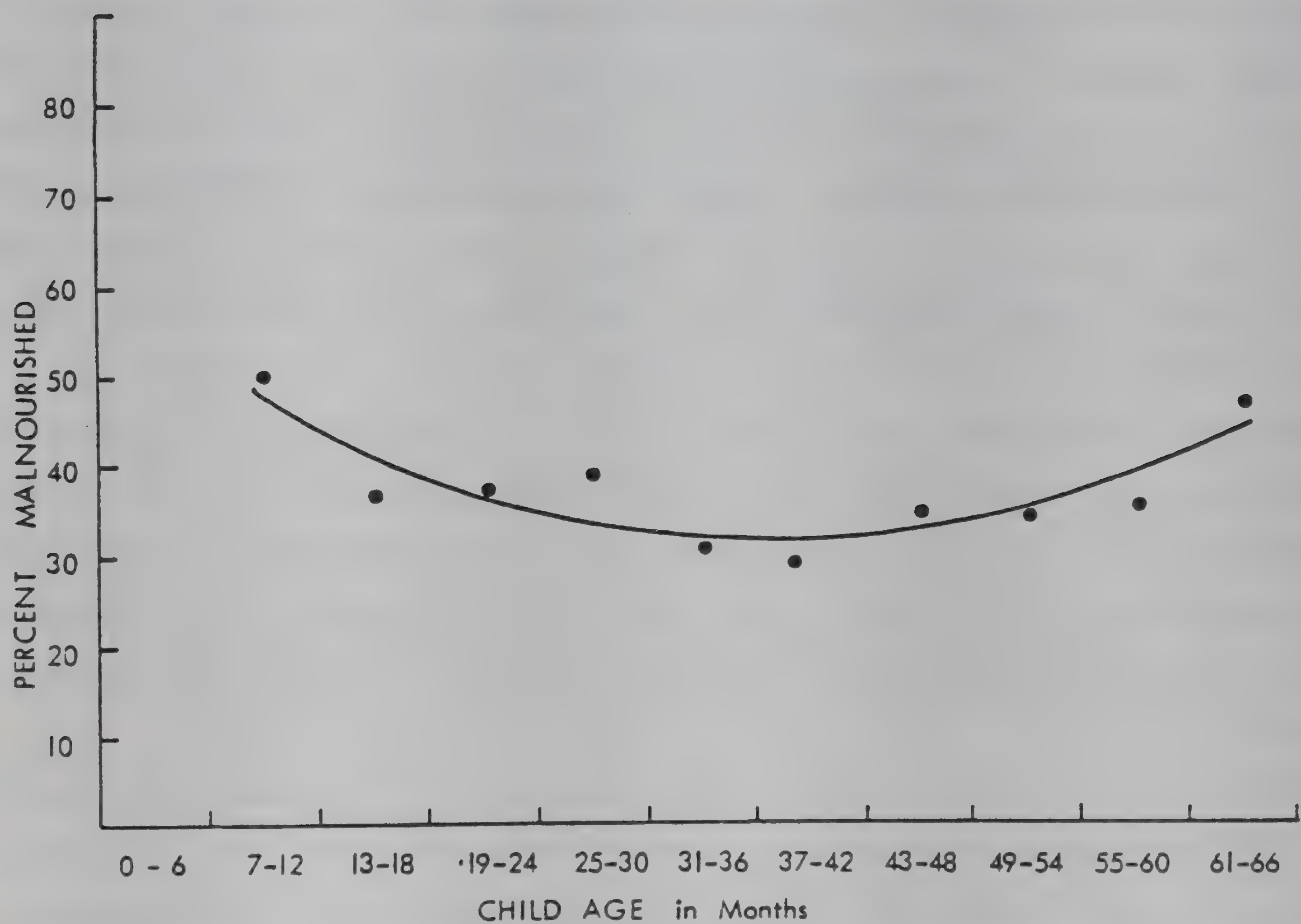


FIGURE 5-2

CHARACTERISTIC CURVE--PAKISTAN-CARE DATA



Malnourished defined as less than 75% of standard used by CARE.

curve is illustrative of the broader question: Can you use the characteristic curve as a simple diagnostic device to learn some things about a community environment? A related question is: Can you learn something about change in nutritional status in a community from examining fluctuations in this characteristic curve over time?

Our answer to both questions is a qualified yes. Examining the relatively few characteristic curves derivable from our data, we observe certain patterns. In Latin American countries, where we suspect the disease-causing agents (especially diarrhea) are critical factors in generating malnourishment, we usually see curves like that found in Honduras. In these countries, although food is scarce, there is sufficient food to allow children to exhibit some natural recovery. However, in the far eastern countries of India and Pakistan, food shortage is the more critical factor causing malnourishment; therefore, the curve takes on a different shape.

Changes in nutritional status over time in a population should show up as shifts in the curve. An overall lowering of the curve shows overall improvement. A lessening of the peak shows an improvement in the worst off. A shifting of the peak to the right or left indicates that a program has had some effect on the causes of malnutrition for the most severely afflicted.

Although the characteristic curve appears to be a useful device for getting a fast appraisal of problems and/or status changes in a community, there are some design decisions that may mitigate its effectiveness if made incorrectly. We have spoken about the problem of setting age in some cultures. Table 5-5 is a histogram of the age variable in the Pakistan data. Note the "local peaks" at twelve, eighteen, twenty-four, and thirty months. This suggests some "rounding" by respondents to the CARE questionnaire—rounding that reduces the reliability of any analysis involving the age variable.

In addition to obtaining the data to make up a curve, there is the problem of categorizing it. The finer the partition of age, the more cases you need. (Note that with about 200 observations in Honduras we used 12-

TABLE 5-5

HISTOGRAM AGE DISTRIBUTION--CARE-PAKISTAN

MIDPOINT	HIST%	COUNT	FOR 8.AGE (EACH X= 1)
8.0000	.1	1	+X
9.0000	.2	2	+XX
10.000	1.0	13	+XXXXXXXXXXXXX
11.000	1.5	18	+XXXXXXXXXXXXXXXXXX
12.000	3.9	48	+XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
13.000	2.0	25	+XXXXXXXXXXXXXXXXXXXXXXXXX
14.000	1.1	14	+XXXXXXXXXXXXXX
15.000	1.6	20	+XXXXXXXXXXXXXXXXXXXX
16.000	2.0	25	+XXXXXXXXXXXXXXXXXXXXXXXXX
17.000	1.2	15	+XXXXXXXXXXXXXXX
18.000	3.5	43	+XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
19.000	2.5	31	+XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
20.000	1.5	19	+XXXXXXXXXXXXXXXXXXX
21.000	2.8	35	+XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
22.000	3.3	41	+XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
23.000	2.9	36	+XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
24.000	3.7	46	+XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
25.000	1.4	17	+XXXXXXXXXXXXXXXXX
26.000	1.3	16	+XXXXXXXXXXXXXXXX
27.000	1.3	16	+XXXXXXXXXXXXXXXX
28.000	1.9	24	+XXXXXXXXXXXXXXXXXXXXXXXX
29.000	2.3	29	+XXXXXXXXXXXXXXXXXXXXXXXXXXXXX
30.000	2.8	35	+XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
31.000	1.9	24	+XXXXXXXXXXXXXXXXXXXXXXXX
32.000	2.3	28	+XXXXXXXXXXXXXXXXXXXXXXXXXXXX
33.000	3.7	46	+XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
34.000	2.4	30	+XXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
35.000	3.3	41	+XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
36.000	3.1	39	+XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
37.000	1.6	20	+XXXXXXXXXXXXXXXXXXXX
38.000	1.5	19	+XXXXXXXXXXXXXXXXXXX
39.000	1.5	19	+XXXXXXXXXXXXXXXXXXX
40.000	1.7	21	+XXXXXXXXXXXXXXXXXXXXX
41.000	2.8	35	+XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
42.000	1.7	21	+XXXXXXXXXXXXXXXXXXXXX
43.000	2.3	29	+XXXXXXXXXXXXXXXXXXXXXXXXXXXXX
44.000	1.2	15	+XXXXXXXXXXXXXXX
45.000	2.5	31	+XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
46.000	2.6	32	+XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
47.000	3.0	37	+XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
48.000	1.9	23	+XXXXXXXXXXXXXXXXXXXXXXX
49.000	1.4	17	+XXXXXXXXXXXXXXXXX
50.000	1.7	21	+XXXXXXXXXXXXXXXXXXXXX
51.000	.8	10	+XXXXXXXXXX
52.000	.9	11	+XXXXXXXXXXX
53.000	1.4	17	+XXXXXXXXXXXXXXXXX
54.000	1.1	14	+XXXXXXXXXXXXXX
55.000	1.9	23	+XXXXXXXXXXXXXXXXXXXXXXX
56.000	1.7	21	+XXXXXXXXXXXXXXXXXXXXX
57.000	1.0	12	+XXXXXXXXXXXX
58.000	.7	9	+XXXXXXXXX
59.000	.2	2	+XX
60.000	.4	5	+XXXXX

3783
NUT100

COMMUNITY HEALTH CELL
LIBRARY AND DOCUMENTATION UNIT

month intervals, and with about 1250 observations in Pakistan we used 6-month intervals.) There is no single rule of thumb on categorization, because the number of observations needed for a given age category is a function of the malnourishment rate itself. If the malnourishment rate is very low, it is necessary to have more observations than if the rate is higher, in order for the numerator to have a sufficient number of points. (Of course, the malnourishment rate is itself variable and subject to design decision; a cut point of 80% of standard raises the rate considerably over a cut point of 75%.) We have found that if there are fewer than fifteen observations in the numerator of the ratio computed to measure malnourishment rate, the curve is unstable—it fluctuates wildly when only a very few children change status.

VARIABLE TRANSFORMATIONS - A NON-LINEAR WORLD

Most models used in analysis are linear ones. Almost all statistics, including analysis of variance and regression, are subsets of the general linear model. Non-linear models are far more difficult to work with and are far less easy to understand. But the world of child malnourishment is comprised primarily of non-linear relationships. The normal growth of children is certainly non-linear, especially in the first eighteen months, and children who are malnourished exhibit an even more non-linear pattern of growth. Most family variables indicating consumption, birth order, family size, etc., exhibit non-linear relationships to nutritional status indicators as well.

The analyst has several alternatives when confronted with this situation. First, he can use linear approximations for non-linear relationships. If the "fit" is not too bad, useful results are still forthcoming. Second, he can attempt to define transformations of non-linear functions which yield linear relations. Finally, he can attempt to construct a basically non-linear model from the outset. It may more accurately portray reality but be subject to greater difficulty when manipulated.

As in most other areas, there are tradeoffs associated with the transformation of non-linear relationships to linear functional form. Each

level of mathematical complexity introduced in the analysis can reduce the intuitive component of the study. In the complex field of child malnourishment, whichever model is used will be at best a partial answer, thereby requiring other judgments when interpreting results. In addition, there is always the possibility of introducing additional errors if the transformation selected is not appropriate. One way of keeping this transformation-introduction-error problem to a minimum is to have a basis in theory for the transformation utilized.

An example of the importance of accounting for non-linearities in nutritional studies is drawn from the Candelaria Revisited data set. In an effort to select the most appropriate functional relationship between nutritional status as measured with a weight/weight(age) measure, we engaged in a revealing empirical investigation. Weight/weight(age) scores were computed for participants and nonparticipants and used as the dependent variable in a regression equation. The explanatory or independent variables were taken to be age and a dummy variable indicating participation. Table 5-6 displays the results of six regressions, each identical except for the transformation of age employed. The overall explanatory power of the regression is highest when the inverse of the square root of age is used. Participation is significant at the .05 level (the program worked) only when the natural log of age was the transformation employed. Thus, the results differ considerably with alternative resolutions to the non-linearity problem.

ATTRIBUTING THE OUTCOME TO THE INTERVENTION--COMPETITORS

We now turn to analytic issues that have far-reaching importance with regard to the philosophy of analysis as well as the mechanics. In the introduction to this report, we made reference to the changing environment and the need to consider competing explanations for observed outcomes. A discussion of the nature of those competing explanations will illustrate their importance. A typology of seven categories of competing explanations is useful.

TABLE 5-6

ILLUSTRATION OF NONLINEARITIES IN NUTRITION ANALYSIS

	R^2	Coefficient	Significance
FORM 1	.140		
Inverse of Age		.312	.000
Participation		-.014	.118
Constant		.928	.000
FORM 2	.140		
Inverse of sq. rt. of Age		.314	.000
Participation		-.017	.056
Constant		.884	.000
FORM 3	.100		
Natural Log of Age		-.046	.000
Participation		-.019	.035
Constant		1.106	.000
FORM 4	.053		
Square Root of Age		-.016	.000
Participation		-.018	.052
Constant		1.046	.000
FORM 5	.025		
Age		-.001	.043
Participation		-.016	.229
Constant		.991	.000
FORM 6	.006		
Age Squared		$-.6 \times 10^{-5}$	.043
Participation		-.012	.229
Constant		.962	.000

Constantly Changing Environment

All community-level nutrition interventions are carried out in environments which are both complex and volatile. The physical environment (climate, water quality and availability, etc.) is often hostile and unpredictable. The socioeconomic environment is typically harsh and is subject to fluctuation due to changes in local, regional, national, and international conditions. The political environment is often highly unstable and, in some instances, repressive.

The intricate workings of this complex environment determine the nutritional status of an individual. At best, an intervention operates on the most exposed edges of this environment as they shape the factors that comprise nutritional status. Clearly, changes in this environment will have strong effects on any outcome measure used to identify changes in nutritional status. Sorting out these effects from program effects is an unbelievable chore.

In almost every intervention considered, some known environmental change can explain observed changes in our outcome measures. In Candelaria, there was steady inflation. In PRIMOPS, a sewer system was added during the middle of the program. In Honduras, there was drought. In KOTTAR, there was drought. In Indonesia, the military disrupted services by shooting at the staff for using a religious structure for a group meeting.

Certain general strategies may be adopted to account for the confounding influences of environmental factors. First, a careful factual summary of the context in which an intervention operates needs to be recorded. This would include chronological records of climatic, physical, social, economic, and political events which occur just before and during the life of the project. Second, reference needs to be made to a general theory of nutritional well being in order to have a checklist of variables to be considered potential sources of perturbation in the nutrition indicators.

Population Aging

The demographic make-up of study populations changes over time. Perceived change in any measure of program impact may, in fact, be due to a change in the age distribution of the population. The relationship between a child's age and his propensity to be at risk was so important that we devoted an entire section to it (see The Necessity of Accounting For Age). However, children are not the only people who age, nor is malnutrition the only important variable related to age.

As an example, in the Candelaria Revisited data there is a clear relationship between fertility and the age of the mother. Table 5-7 shows a steady decline in the number of women who became pregnant between 1974 and 1976 as the age of the women increases. The Chi-square statistic is an indication of the statistical significance of this relationship. If one were studying changes in fertility in a given population, those changes would have to be computed as the difference between observed change and expected change due to aging. In our comparison of participants and nonparticipants with respect to fertility, we had to control for the age differences in the two groups.

The implication of recognizing the importance of population aging in the design of an analysis of an intervention is that sufficient data must be collected to enable the analyst to control for age. This may result in a substantial increase in data acquisition costs.

Project Aging

We have noted that nutrition interventions go through a life cycle. During the early stages of the project the members are devoting more energy to entrepreneurial efforts and less to making the project run smoothly. Later emphasis may shift to operations, such as more carefully training a staff, better followup procedures, improved information handling, and gaining of acceptance of the program in the community. Finally, there are the mature stages where much of the excitement and newness has worn off, and possibly the external resources have either been held constant or tapered off in real terms due to inflation. Often, persons who

TABLE 5-7

THE RELATIONSHIP BETWEEN AGE AND FERTILITY--CANDELARIA REVISITED

Pregnant since 1974	AGE						
	Younger than 20	20-24	25-29	30-34	35-39	40 and over	Total
Yes	20 (14)	104 (78)	72 (66)	39 (53)	37 (39)	11 (34)	283
No	5 (11)	30 (56)	41 (47)	52 (38)	30 (28)	47 (24)	205

CHI-SQUARE = 72.69: SIGNIFICANT AT ANY LEVEL

were at first volunteers now expect remuneration, further stretching a limited budget.

To compare a new project with an old one could therefore by very misleading, but more typically, to compare a new component with an old one in the same project and not take these factors into account could lead to misleading conclusions simply due to being at a different stage of the project life cycle.

We are unable to point to specific examples where life cycle changes explain observed program outcomes; however, we can speculate that the declining rate of improvement over time in Candelaria may reflect project aging. Similarly, the lack of dramatic change in PRIMOPS may be, in part, due to the fact that our first round of data was really the program's fourth. (It was the first after an 18 month hiatus.)

Participation Biases

Nutrition programs are almost always in the process of expansion or contraction. There is natural expansion and contraction due to births and aging in the participating families. Also, there is migration into and out of the study area. Finally, the study area itself may be growing or shrinking. Changes in the composition of the target population over time as the participants change create bias in subsequent analysis.

To illustrate this further, we will consider two known sources of participation bias. Programs just starting out often begin serving those families closest to the facility from which services emanate. Over time, services are extended to more geographically distant areas. Often, because central facilities are located in urbanized areas, this movement to more distant areas is a movement between people of different status and/or risk level. Changes in impact must be measured while controlling for changes in the population. In KOTTAR, Honduras, and Indonesia, program participation was gradually expanded over time to villages more distant from the epicenter.

The other source of participation bias is the self-selection process of "dropping out." Participants drop out of programs because they no longer

need the services, because they are no longer able to reach the service center, or because they do not perceive that they are being helped. In each case, a bias is generated in the analysis of data from two or more different time periods. In PRIMOPS, the dropout rate of participants increased as nutritional status deteriorated (Table 5-8).

Improvement Due to Natural Phenomena

Thus far, we have identified several known sources of distortion of results of analysis and, in essence, have argued that program outcome (change in nutritional status) must subtract those components of outcome attributable to the sources of distortion. One other important source of change in an outcome measure is the natural change that would occur in the population if it were left alone. As analysis focuses more and more on the malnourished, the spontaneous or natural improvement of some children contributes increasingly to observed outcome changes. In the absence of an intervention, some grade III malnourished children will recover. If one just looks at grade III children, one can only attribute to the intervention any recovery above and beyond the natural recovery rate.

This concept is very important for the design of an analysis strategy. It suggests that one should not zero in on just the worst cases for data collection and analysis, unless there exists an a priori quantification of the improvement due to natural phenomena.

Variations in Reliability Over Time

Data collected over time are subject to variations in accuracy. These variations in accuracy, if continually in one direction, may alter the results of any analysis of program outcomes.

We have some evidence that variations in the reliability of data gathering exist. Our evidence pertains to variations occurring in recording and keypunching. Although we have no example indicating that measurement accuracy deteriorates, we know that scales loose accuracy, and can hypothesize that people repeatedly using those scales lose accuracy too. Table 5-9 shows the deterioration of the PRIMOPS data over time.

TABLE 5-8

DROP OUT RATE - PRIMOPS

Nutrition Status Round I	Number	Percentage Missing In Round 2
N	3494	.1
1	1047	.2
2	479	.2
3	93	.3

TABLE 5-9

RELIABILITY VARIATIONS IN PRIMOPS

	Original	Range Errors	Duplicates	Total	Pct.
Subset 1	3045	150	62	212	6.9
Subset 2	6253	982	124	1106	17.6
Subset 3	8020	1285	473	1758	21.9
Subset 4	4383	810	88	898	20.4
Subset 5	2918	455	262	717	24.5

We identify five distinct subsets of data—each collected in a later time period than the previous subset. The errors are the number of records culled from the original due to range violations (illegal codes) or duplicate entries against a single ID. The increase in error rate is apparent.

Hawthorne Effects

When attempting to attribute changes in nutritional status to program characteristics, it is easy to assume that these changes are due to the visible and tangible aspects of the intervention. However, it is always possible that there may be some process-related characteristic of the project that in fact has contributed to the outcome. We know that in some circumstances simply the visibility of the participants in a program will cause changes in behavior to occur. A related example occurred during the Candelaria experiment, when the mothers of families participating in a health and hygiene education component of the project cleaned up their homes in order to look "proper" when the doctor visited their homes. While the hygiene education program was important, the principal motivating factor was the expected visit from the doctor.

The seven "competitor" categories embrace the source of many of the confounding factors which enter into analysis. The additional source of variation in any outcome measure is the random or stochastic effect. A statistical model builder always acknowledges that even if every relevant variable could be included in his/her model, some variation in the outcome variable will still be unexplained. This is the stochastic effect. The use of statistics is really the application of tools to determine if the values taken on by the outcome variable could be attributed to this stochastic effect alone, given the hypothesis that the other variables have no effect. Thus, the use of statistics is, in fact, the control for the stochastic confounding effect.

ATTRIBUTING THE OUTCOME TO THE INTERVENTION—CONTROLS

The method most often espoused to account for confounding effects (potential competing explanations) is the use of controlled experimentation.

The idea behind the use of controls is that if two groups of children with similar characteristics (no statistically significant differences on important variables) differ only with regard to their participation in the program, a comparison of changes in the outcome variable over time in the two groups will measure only that component of change attributable to the program. The assumption, of course, is that the two groups will be subject to the same environmental fluctuation, same degree of aging, same types of participation bias, same degree of spontaneous recovery, and same reliability variations. (Both the Hawthorne Effect and the Project Aging factors are important in partitioning the overall program impact into components.)

We can identify three types of controlled experimentation. The first utilizes matched pairs in the "treatment" group (participants) and the control group (nonparticipants). Matching in this way is generally regarded as immoral by practitioners of interventions. The very definition of control implies the withholding of services from one group. In any real setting, it would be immoral for a human being to refrain from treating a child who had been identified as one in need of specific attention, and where it was within the power of the individual to provide that attention.

More common outside of a laboratory are quasi-experimental approaches where groups believed to be similar overall are selected—one for treatment and one as control—but are not comparable individual by individual. Use of a nonparticipating but geographically proximate village as a control is quasi-experimental. Here again one could raise the morality issue; however, in either a circumstance of limited resources or one of a phased introduction of a service, the concept of control becomes acceptable once again. But there are other problems. Communities in the developing world are in highly volatile environments subject to wide fluctuations over time. If there is too much physical distance between communities, it is almost impossible to assume that their environments change in parallel fashion over time. If they are too close together, it is difficult to keep the services from spilling over into the control group.

In the interventions studied in depth as part of our project, "controlled"

designs were attempted only three times. One is illustrative of the distance between communities. In Candelaria Revisited, participants in the program were compared to nonparticipants living in Candelaria two years after the program was terminated. For the most part, nonparticipants were in-migrants. Statistical tests showed major differences in the populations—primarily with regard to the age distributions of both children and mothers.

The spillover effect is illustrated by the Esperanca project. The control village for Mojui dos Campos was a suburb of that village. People from that suburb used the health services available in Mojui dos Campos and, therefore, were exposed to the Esperanca health post. We can conclude that although quasi-controls may be possible, they are, at best, difficult to establish. A research design dependent on the success of establishing a control has so high a probability of failure that it should not even be attempted.

The third type of control is the use of the treatment group as its own control. Such variables as duration of exposure and/or intensity of exposure can be used to compare subsets of the population which are similar to each other in most respects other than time in program or intensity of participation. This technique was applied using the original Candelaria data set with success. The study showed that long-term participants were better off than short-term participants of the same age (Drake, Fajardo; 1976). In Thailand, the technique was applied to show no program effect, using an intensity-of-exposure measure (frequency of consumption of fortified rice) as the control (Gershoff et al., 1977). There are both strengths and weaknesses to this technique. The strongest point in its favor is that, in any practical setting, it is relatively easy to gather data on the duration of participation along with other outcome and process variables. There is no need to query nonparticipants, since the first observation of the participant in the intervention coincides with no time in the program, and subsequent observations record correspondingly increasing exposure to the intervention. The weakness in this technique is that competitive explanations based on environmental changes, population aging,

project aging, reliability variation, and participation biases may still confound the outcome. Consequently, we have combined this type of control with the use of statistical controls to resolve these difficulties to the extent possible. Of all confounding effects, the Population Aging one is the most important. A resolution of the problem is the use of the characteristic curve which relates population malnourishment as a function of child age. A comparison among a family of characteristic curves at varying durations of an exposure to the intervention can control the effect of this confounding variable. Changing environmental conditions, project aging, and participation biases can be dealt with statistically. Reliability variations need to be checked further.

In the last paragraph, we mentioned statistical controls. Although we might call this a fourth type of controlled experiment, we choose to view it as an analytic technique--not an experimental design. Statistically derived controls are accomplished by selecting subsets of the population, either within the intervention or outside it, that are matched with the participants on important dimensions. For instance, when we were looking at the long-term effects of the Candelaria Promotora program, a comparison was made between families who had participated in a prior time period with families who had recently migrated to Candelaria and therefore never participated. When it was found that the immigrant families were typically younger, a subset of immigrants who "matched" the participants in age was selected for comparison purposes. This "matching" process relied upon a priori knowledge of the processes involved in order to know which among the myriad of variables are important enough to be used for comparison. In practice, to go beyond a few critical variables is unwise and impractical. Other multivariate statistical models serve as techniques for establishing statistical controls such as regression analysis, illustrated in the last section of this chapter.

In summary, the use of strictly defined controls in the analysis of community-level nutrition programs is usually ineffective or unmanageable. For time series data sets, use of the treatment group as its own control and the setting up of appropriate statistical

controls to account for anticipated confounding results is the best strategy.

OUR OVERALL APPROACH TO ANALYSIS

In both the opening chapter and the early sections of this chapter, we built a foundation for an approach to analysis of community-level nutrition programs. Most of the elements of that foundation stand alone as important considerations in any analysis of such programs (for example, the need to explicitly account for age of child in all analyses involving nutritional status). Now, we will attempt to forge those elements into an overall approach to analysis—one that we apply ourselves and recommend for others.

Our approach is heavily dependent on three of our earlier conclusions—so heavily that they bear repeating here.

A. Analysis must be participatory. The correct interpretation of results requires the input of program managers and staff whose knowledge of the population and environment is far greater than that held by any outside evaluator.

B. Analysis should consist of the formulation and subsequent testing of competitive explanations of observed outcomes. The "competing explanations" must be fully articulated and subjected to intense scrutiny before any single interpretation or set of interpretations is accepted as "truth."

C. Controlled experiments in the real world (social settings) are extremely difficult to effectuate. A research design counting on a comparison with a control has a high probability of failure and should be avoided. Other designs, though not as conclusive when executed, have a higher probability of success and are more likely to yield some usable results.

Given these three conclusions, a specific approach to analysis becomes apparent:

1. Formulate a logical construct of all possible, significant competing explanations of outcomes derived from the data.

2. Perform some simple analyses to gain insight into the

underlying phenomena occurring at the community level.

3. As appropriate, use more sophisticated analytic methodologies and/or other known facts about local events to narrow the list of "competitors."

4. Present the results of the analysis to project personnel to learn their views on remaining competitors (and to gather any additions to the list they may provide).

5. Iterate steps 3 and 4 until competing explanations of outcomes are as few as possible.

6. Incorporate conclusions into the local decision process, thereby surfacing areas which need further analysis and, hopefully, improving the intervention as project leaders respond to the findings of the analysis.

In the following paragraphs, we will attempt to elucidate each of these steps and speak to the issues associated with "making them happen."

Enumerate Competitors

Implicit in the notion of formulating a logical construct is the existence of a theory or model of the nutrition system upon which to base competitive explanations of observed outcomes. It is surprising how often such a theory or model is missing; that is, neither the project personnel nor the analysis team can offer a detailed explanation of the "cause and effect" relationships underlying their own activities.

The seven categories of competitors described earlier suggest many components of such an underlying model. Most important of all is the recognition that the changing environment may account for changes in outcome variables far greater than those anticipated from a given intervention strategy, and that these environmental changes must be "controlled" for in the analysis.

We would like to illustrate what we mean by underlying theory or model by describing one model used in this fashion by Community Systems Foundation (CSF): the Nutrient Flow Model. This model was designed in Cali, Colombia, through a collaborative effort of scientists and doctors from the Universidad del Valle and CSF. The model was intended to guide

in the formation of a strategy of intervention as well as the analysis of its impact. That is, the model was intended to provide a theoretical basis for diagnosing the causes of malnutrition in a community, determining what types of interventions are most likely to alleviate those causes and, finally, prescribing a data collection system and mode of analysis for verifying the expected impact of the proposed interventions.

The model is based on the "nutrient flow theory." The primary concept in this theory is that the level of malnutrition in the community, household, and/or individual is related to the difference between nutrient requirement and nutrient availability--a difference called the "nutrient gap." Implementation of this theory calls for the elaboration of the determinants of nutrient availability and nutrient requirements.

Nutrient availability is determined by the complex process by which food is produced, distributed, and consumed. The measurement of nutrient availability, therefore, requires that food flows be measured and that losses occurring at the various stages be determined; for example, losses during harvest, storage, transport, processing, unequal allocation among families and non-optimal allocation within families, and lastly, in preparation.

Nutrient requirements are determined by several individual characteristics such as weight, age, activity levels, and sex. Of primary importance in developing countries is the ability of an individual to absorb nutrients--an ability that is adversely affected by a complex of diseases and disease-causing factors.

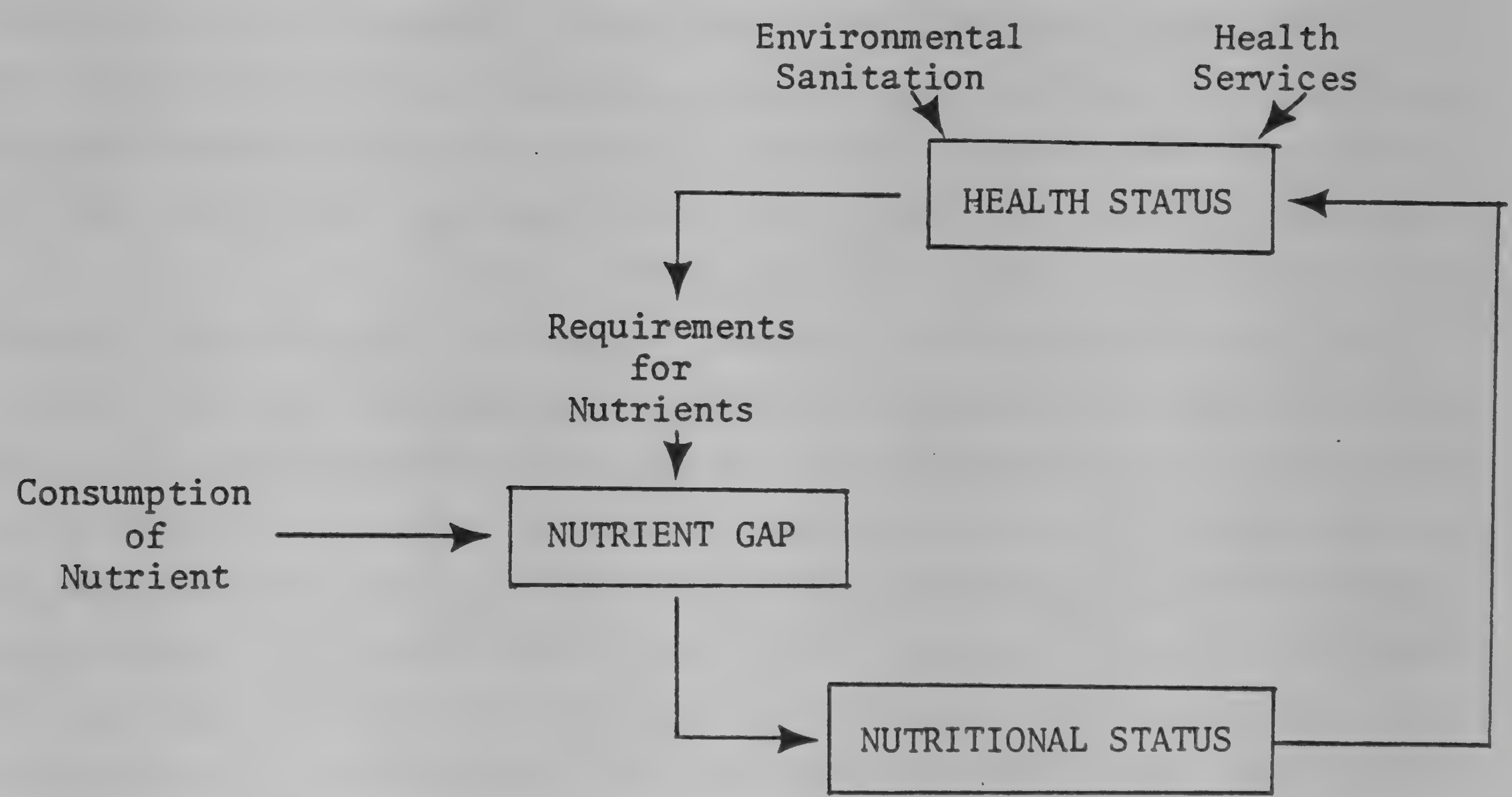
The model is applied as a diagnostic device by measuring the contributing factors to both the nutrient availability and the nutrient requirement. The measured size of the gaps at the individual, household, and community level, the nature of nutrient losses in the system, and the kinds and magnitudes of disease-causing factors suggest where interventions might produce the greatest impact. Conceptually, the model looks quite simple (Figure 5-3).

The simplicity should not mask its profound implications:

(a) The model illustrates the negative consequences of its "positive feedback" loop; that is, a decline in health

FIGURE 5-3

THE NUTRIENT FLOW MODEL



status leads to a growing nutrient gap which adversely affects nutritional status which results in a further decline in health status.

(b) The dynamics of this and other feedback loops are not well understood and are hardly quantifiable; for example, no one knows the magnitude of shift in nutrient requirement resulting from the incidence of diarrhea disease.

(c) Program objectives can be stated in terms of anticipated increases in consumption or reductions in requirements, and can be evaluated by examining measurements of those quantities.

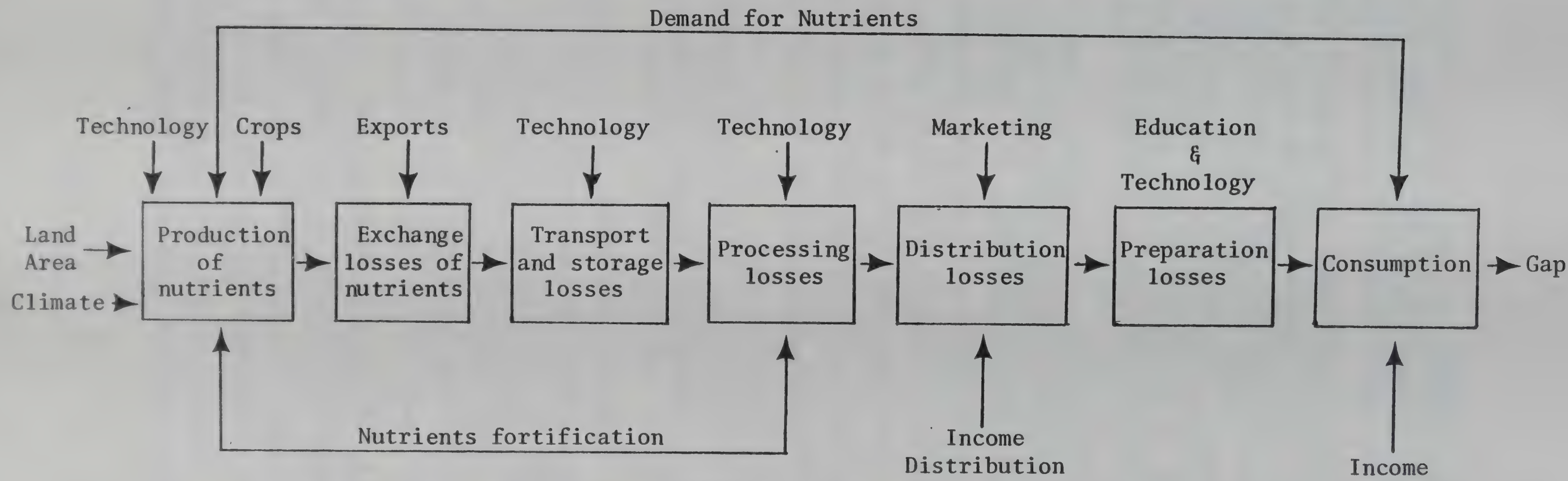
To "flesh out" the model, one must create a "consumption section," that is, a procedure for measuring consumption must be conceptualized and effectuated. The original version of the model used a procedure based on the national food-balance sheets which have been used for several years by FAO and the United Nations (see Figure 5-4).

The implementation of an analysis (diagnosis, intervention, impact analysis) using the Nutrient Flow Model is excessively difficult. It is a "data hungry" model, requiring vast amounts of resources going into the measurement of consumption patterns. Yet even this model is an oversimplification, particularly in its treatment of the determinants of consumption such as income. As a result of our project, we would argue that income, treated as a small part of Figure 5-4 (almost an afterthought), is one of the most important determinants of consumption, and that the causes of poverty (land tenure issues, unemployment, political repression, etc.) may be more important targets for intervention than those suggested by the food-balance sheet approach taken in the model as shown.

In summary, we argue that some model or theory is a necessary entity to guide in the selection of a set of competitive explanations for observed outcomes. Furthermore, we argue that particular qualities of the local situation dictate which components of a general "nutrition system" model are most important in that given situation.

FIGURE 5-4

THE NUTRIENT FLOW MODEL - CONSUMPTION COMPONENT



Perform Simple Analyses

In our original presentation of the six steps of our analysis approach, we accentuated the word "simple." We believe strongly that simplicity in analysis is a virtue. It may seem strange to the reader that immediately following a plea that a complex model be used as a basis for identifying competing explanations for outcomes, we turn around and make a plea for simplicity of analysis. The two are not incompatible. Complexity of the model contributes to the richness of the competing explanations; simplicity of analysis contributes to the ease of interpretation of tests of the validity of a particular explanation.

Our strong advocacy for simplicity comes, in part, because of our commitment to participatory analysis. It is difficult enough for highly trained social scientists to grapple with the complexities of advanced analytic techniques; it is almost ludicrous to expect project staff in the field to do so. Thus, we advocate the use of simple graphs to show the relationships between pairs of variables and the use of two-way contingency tables (and simple Chi-square tests) prior to employing multivariate techniques.

With regard to simplicity in analyzing nutrition interventions using anthropometrics, we argued that the characteristic curve (plot of percent malnourished against age categories of children) is a useful simplifying device. Often, competing explanations for outcomes can be tested by comparing characteristic curves for subsets of the population. For example, the difference (if there is one) between new program participants and continuing program participants with respect to nutritional status can be seen easily by plotting the curves for both groups.

Another useful device is the transition matrix, which illustrates neatly the movement of children between categories of malnourishment. The matrix is an immediate indication of the effect of the intervention on the children who are worst off. To illustrate the simple lessons to be learned from the characteristic curve and transition matrix, we use the PRIMOPS data. A local "normal line" was used to determine the weight/weight(age) ratios, and categories are formed as follows: less than sixty percent,

grade III; sixty percent up to seventy-five percent, grade II; seventy-five percent to eighty-four percent, grade I; and eighty-four percent and up, normal.

The characteristic curves for the three rounds for which we have data are presented in Figure 5-5. The three curves show a slight improvement in the nutritional status of the population over time.

The transition matrix between round 1 and round 2 is presented in Table 5-10. The diagonal elements in the matrix indicate the number of children who remained in the same category for the two rounds. The elements below the diagonal indicate the number of children who improved, while the numbers above indicate the numbers whose nutritional status dropped. We see that the nutritional status of 293 children deteriorated while 284 children improved. Ignoring the missing columns for a moment, we see that over half the grade III children in round 1 improved (a much greater number than in any other category) but that an even greater number entered the category. Another interesting fact that can be seen in the matrix concerns the rate of disappearance from the categories. We see that 57 of 89 (64%) grade III children in round 1 were not observed in round 2. In the other categories, 45%, 42%, and 38% of the children were not observed. Grade III children dropped out at the highest rate.

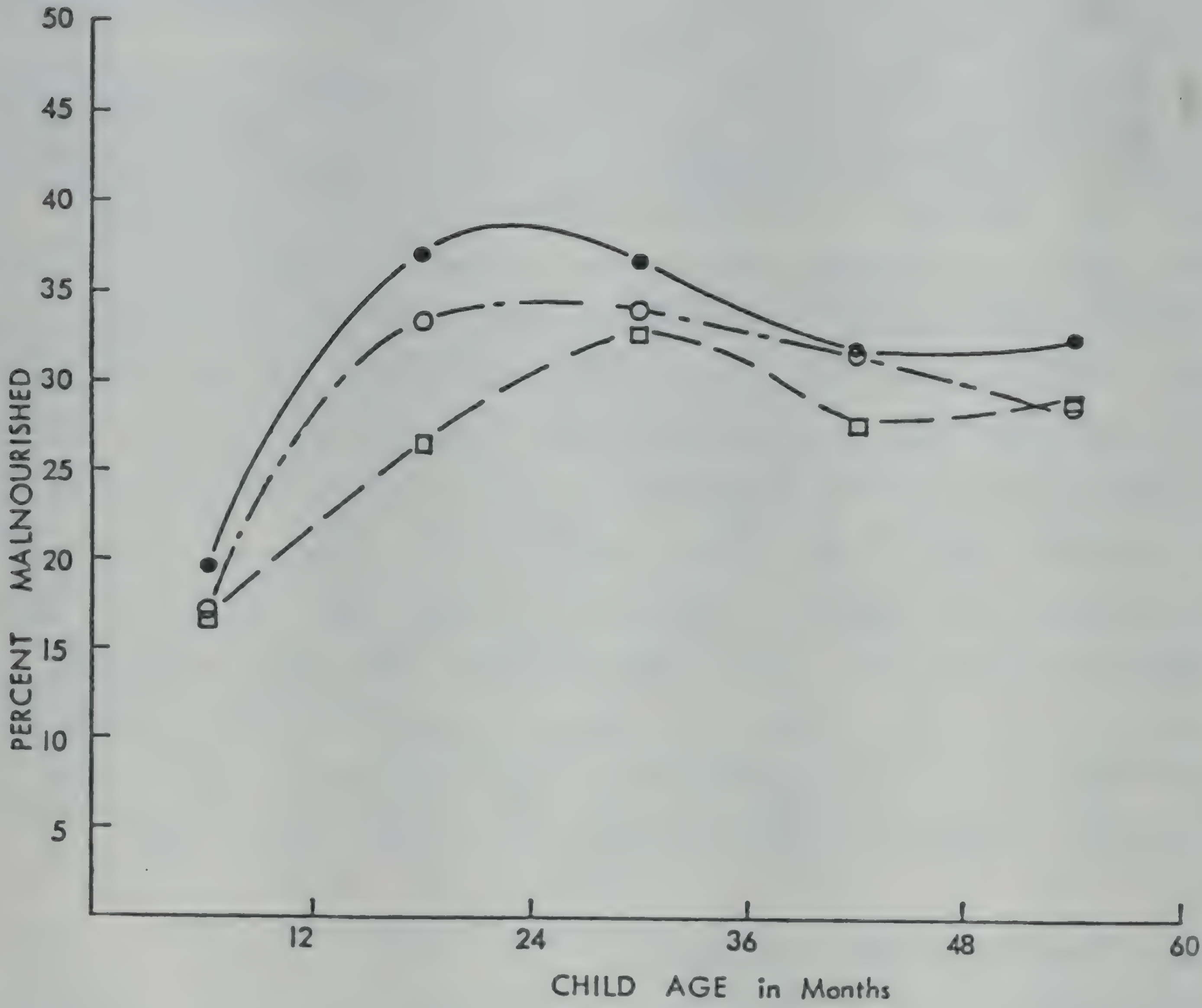
Use More Sophisticated Analyses

Unfortunately, the examination of the validity of complex explanations of outcomes often requires complex analysis. Multivariate techniques are often required to sort out combined effects of various factors on nutritional status or any other outcome indicator.

Parametric Tests. To illustrate the range of statistical techniques available for more sophisticated analysis, we will describe the settings under which some are applicable. Almost all statistical analysis utilizes the general concept of hypothesis testing. In short, a "null" hypothesis (often a hypothesis that no relationship exists between a set of variables) is posited, as is an "alternative" hypothesis (often a hypothesis that the

FIGURE 5-5

CHARACTERISTIC CURVES--PRIMOPS--CALI, COLOMBIA



DATA: Percent Malnourished by Rounds (1976 - 1978)
1 to 5 year old children, one year age categories

Age Categories	1	2	3	4	5
Round 1	19.9%	37,1%	36.8%	31.9%	32.5%
Round 2	17.1%	33.4%	34.1%	31.3%	28.2%
Round 3	16.7%	26.6%	34.0%	27.4%	29.0%

TABLE 5-10

TRANSITION MATRIX - ROUND 1
TO ROUND 2 - PRIMOPS

From Category in Round 1	To Category in Round 2					
	N	I	II	III	MISS	TOTAL
N	1940	193	21	1	1355	3510
I	171	361	59	3	434	1028
II	16	79	150	16	216	477
III	1	1	16	14	57	89
MISS	2897	712	335	56	--	4000
TOTAL	5025	1346	581	90	2062	9104

relationship exists and usually stating something about the nature of that relationship). The test is a computation of the probability (under certain assumptions) that the observed data could appear, given that the null hypothesis is true and that there is a random disturbance (an element of chance) in the world. If the computed probability is sufficiently small, we reject the null hypothesis and act as if our alternative hypothesis is true.

Difference of Means Test. Simple analyses include such tests as Student's Differences of Means Test. For example, suppose we have measured nutritional status of two groups of children--children who participated in a program and children who did not. For a child of a given age, nutritional status was determined by computing the ratio weight/weight(age). A reasonable hypothesis for comparing the two groups is that there is no difference between the means of the nutritional status measurements in each group. Alternatively, there might in fact be a difference.

One must be cognizant at all times of the assumptions underlying the use of any statistical test. (Using computers, it is too easy to generate all kinds of numbers without fully understanding their meaning.) A difference of means test assumes that the populations (groups) are distributed normally (with respect to the variable of interest) and that the population variances are the same. If the sample size is large enough, there are ways to test for differences between means even when the variances are unequal.

For the balance of this chapter, we will draw examples from our Candelaria Revisited data set to illustrate each test. The tables presented are the actual computer-generated tables. To assist the reader in examining the tables, Table 5-11 defines the variables used; unfortunately, computers use abbreviations which are often difficult to interpret.

Table 5-12 is the computer output from a differences of means test to see if program participants differed from nonparticipants with respect to nutritional status. The MIDAS program used to generate this table prints out the significance level of the test—that is, the minimum probability for

TABLE 5-11

VARIABLE LIST FOR SAMPLE OUTPUT

Variable Number	Variable Name	Description
443	PCW 76	Nutritional status of a child as measured by computing weight as the percentage of a reference population standard (the Candelaria standard) weight for the child's age (weight/weight(age))
500	PROGPART	Indicator of the participation status of the child's mother: 1 = yes, 2 = no
502	PROGPART	Indicator of the participation status of the child: 1 = child over 2, participant; 2 = child over 2, nonparticipant; 3 = child under two, mother participated; 4 = child under two, mother did not participate
11	CMLTIS	Number of other children in the family less than 15 years old
406	Age	Age of the child in months
706	Age Sq.	The square of the age of the child in months
839	INC. SUB.	Monthly income of the family added to any monthly subsidy received by the family

TABLE 5-12

A SAMPLE DIFFERENCE OF MEANS TEST;
CANDELARIA REVISITED

TWO-SAMPLE T-TESTS

VARIABLE	PROGRAMT	YES	NO	TEST STATISTIC	DF	SIGNIF
443.	MEAN	93.708	94.382	T=-.70439	846	.4814
POW75	VAR	160.75	228.86	F= 1.4237	360, 486	.0001
(TOTAL= 859)	N	487	361	PROB(1ST MEAN<2ND |DATA)=		.7534

which the null hypothesis can be rejected. In this case, if we select a level of significance higher than .48, we can reject the null hypothesis. Traditionally used levels of significance are .01 or .05; therefore, we conclude that the means are **not** statistically different.

Note, however, that the significance of the F-test for differences of variance is sufficiently small to conclude that the variances are not the same in the two groups. The printed number is a one-tailed test and must be doubled; however, for any significance level higher than .0002 (including the usual .01 or .05), we can assert that the variances are statistically different. This renders our application of the t-test suspect. In this instance, our sample sizes are sufficiently large to permit us to use the Bayesian posterior probability provided by MIDAS (the PROB 1st MEAN > 2nd/DATA) to compute a level of significance concerning the difference between means. A discussion of this test is beyond the scope of this section of the report; the results, however, are almost identical. The difference in means is not statistically significant.

Analysis of Variance. More complex analysis is needed when we believe that there may be "confounding" variables; that is, observed differences (or masked differences) may be explained in terms of other important factors describing the population. For example, in measuring differences (non-differences) in nutritional status as described above, age distributions of the groups may explain observed outcomes. We have used three related techniques for dealing with the multi-variate case: analysis of variance, regression analysis, and analysis of covariance. All three techniques are special cases of a "general linear model." In fact, regression enthusiasts view all three techniques as one in the same; both analysis of variance and analysis of covariance are special cases of regression. Although we too fall into this class of analysts, we present our discussion in the more traditional way.

Analysis of variance is applicable when the suspected confounding variables are "qualitative" (nominal or ordinal). Regression is applicable when the confounding variables are "quantitative" (ratio or interval).

Analysis of the covariance is used when both qualitative and quantitative variables are believed to confound a hypothesized relationship. In all three cases, we speak of a dependent variable and explanatory or independent variables which are believed to explain the variance in the dependent variable. (The qualitative variables can be entered into a regression model by using dummy variables; hence, the regression enthusiast.)

In terms of our example, the explanation of differences in participants and nonparticipants, we have already alluded to other variables which are related to nutritional status—for example, age. A particularly interesting confounding situation arises in our Candelaria Revisited data. The participation of a child is not clearly defined for children less than two years of age. The program in Candelaria terminated in 1974; the survey was done in 1976. Children under two could not have participated directly in the program. In our initial example, the mother's status was used to define participation. Children under two whose mothers had been in the program were called participants. A logical question which might be asked is the difference between actual participants over two, nonparticipants over two, chidren under two whose mothers had participated, and children under two whose mothers had not participated.

This is a comparison of means from four groups—a case not covered by the simple t-test for differences of means. This is a one-way analysis of variance. (The one-way refers to the existence of only one qualitative variable.) Pairwise comparisons between the four groups could be done in six separate tests, but such an approach would not yield an overall conclusion about differences. Also, the overall level of significance of the separate tests is **not** the same as the true level of significance computed in an analysis of variance.

Table 5-13 is the computer output from a one-way analysis of variance to test for differences of means in the four groups of children as defined by our revised definition of participation. The level of significance of the overall analysis of variance is so small (we can reject the null hypothesis for all values greater than that reported) that we can easily conclude that the four groups do differ significantly with respect to nutritional status.

TABLE 5-13

A SAMPLE ONE-WAY ANALYSIS OF VARIANCE;
CANDELARIA REVISITED

```
UNIVARIATE 1-WAY ANOVA


ANALYSIS OF VARIANCE OF 443.PCW76  N= 848 OUT OF 859

SOURCE                  DF SUM OF SQRS  MEAN SQR  F-STATISTIC SIGNIF

BETWEEN                  3    6015.8     2005.3      10.948     .0000
WITHIN                 844    .15459 +6  183.17
TOTAL                  847    .16061 +6  (RANDOM EFFECTS STATISTICS)

ETA= .1935  ETA-SQR= .0375  (VAR COMP= 9.0448  %VAR AMONG= 4.71)

EQUALITY OF VARIANCES:  DF= 3, .89234 +6    F= 21.101     .0000


PROGRAPT      N     MEAN     VARIANCE   STD DEV

GT2CP       308   93.552     119.26     10.921
GT2NP       276   91.065     150.29     12.259
LT2FP       113   96.566     291.89     17.085
LT2NP       151   98.331     293.05     17.119

GRAND       848   93.995     189.62     13.770
```

Worthy of note are the magnitudes of the means. Our simple t-test showed the participants worse off (but not statistically significant) than the nonparticipants. In Table 5-12 we see why. There were a larger number of children under two whose mothers did not participate than whose mothers did and those children, as a group, were "best off." Among children over two, the participants were actually better off than nonparticipants. So we see, the more complex analysis does tell us things we could not detect from the simple analysis.

Regression Analysis. Regression analysis is used when the independent variables are quantitative. Special forms of regression analysis can be used to introduce qualitative variables in certain situations. Regression analysis is particularly useful because it not only allows the analyst to test for the existence of a relationship between an independent and dependent variable, but also allows the analyst to quantify that relationship. This technique provides the analyst with a coefficient relating each independent variable with the dependent variable, a coefficient that estimates the amount of change that takes place in the dependent variable for each unit change in the independent variable.

Table 5-14 is an example of a regression analysis using the same Candelaria Revisited data set. We have referred to the importance of age in explaining nutritional status, and therefore we include age in our equation. Furthermore, because of the growth spurt in very young children and the natural improvement of most children as they grow older, we include the square of age in the equation to account for the known non-linear relationship between nutritional status and age. To illustrate the multivariate capabilities of the technique, we include two other variables: number of other children in the family less than fifteen, and monthly income plus monthly subsidies granted to the family. Note that we do not yet deal with participation; it is not a quantitative variable.

The critical values in the typical multiple regression printout are the R^2 (R-SQR on the MIDAS printout), the coefficients, and the significance of the individual coefficients. Our R^2 of .16927 says that the set of

TABLE 5-14

A SAMPLE REGRESSION ANALYSIS;
CANDELARIA REVISITED

LEAST SQUARES REGRESSION

ANALYSIS OF VARIANCE OF 443.PCW76 N= 733 OUT OF 859

SOURCE	DF	SUM SQRS	MEAN SQR	F-STAT	SIGNIF
REGRESSION	4	22173.	5543.2	37.083	.0000
ERROR	728	.10882 +6	149.43		
TOTAL	732	.13099 +6			

MULT R= .41142 R-SQR= .16927 SE= 12.226

VARIABLE	PARTIAL	COEFF	STD ERROR	T-STAT	SIGNIF
CONSTANT		107.86	1.6358	65.934	0.
11.CHLT15	-.19678	-1.2427	.22948	-5.4152	.0000
406.AGE	-.35516	-.88509	.86341 -1	-10.251	.0000
706.AGESQ	.33617	.11295 -1	.11728 -2	9.6308	.0000
839.INC.SUB	.13585	.39731 -2	.10739 -2	3.6996	.0002

independent variables used explains 17% of the variance in the dependent variable. Although this might seem low, in social science research, values near 20% are common; values near 30% often cause euphoria in the analyst. The significance of the coefficients is evident at just about any level, and the coefficients themselves give a quantitative measure of the relationship; for example, for each child under 15 in the family, nutritional status drops by 1.2 percentage points.

Analysis of Covariance. Analysis of covariance is most often used in the context of laboratory experiments where strict control groups can be set up. In this case, primary interest is in the effect of a treatment on a dependent variable; however, other quantitative variables are believed to affect that same dependent variable as well. In the non-laboratory world of the social sciences, analysis of covariance is used when a functional form is hypothesized where one or more of the independent variables are qualitative but all the independent variables are of **equal importance** in the analysis.

Our example from Candelaria Revisited adds the variable participation into our hypothesized relationship (equation). We are now interested in whether the regression lines that can be calculated for each "group" (value of the participation variables—yes or no) are the same or different. Also, if they are different, is the difference a change in the intercept (constant) or a change in the coefficients (rates of change of the dependent variable with the independent variables)?

Table 5-15 is a one-way analysis of covariance generated by the MIDAS computer program. This table shows four significance levels for four different tests in its topmost segment. The first is a test for the significance of the covariates (quantitative variables) in explaining the variance of the dependent variable above and beyond the variance explained by the qualitative variable (participation). As expected, the null hypothesis is rejected; the co-variates have a significant effect on nutritional status. The second test is the reverse; that is, it is a test for the significance of the qualitative variable in explaining variance in the

TABLE 5-15

A SAMPLE ANALYSIS OF COVARIANCE; CANDELARIA REVISITED

1-WAY COVARIANCE MODEL

ANALYSIS OF VARIANCE OF 443.PCW76 N= 733 OUT OF 859

SOURCE	DF	SUM SQRS	MEAN SQR	F-STAT	SIGNIF
BETWEEN MEANS	1	118.94			
COVARIATES	4	22247.	5561.8	37.223	.0000
ERROR	727	.10863 +6	149.42		
- - - - - - - - -					
REGRESSION	4	22173.			
EQUAL ADJ MEANS	1	193.43	193.43	1.2945	.2556
ERROR	727	.10863 +6	149.42		
- - - - - - - - -					
OVERALL REGRESSION	4	22173.			
EQUAL REGRESSIONS	5	746.42	149.28	.99868	.4175
EQUAL ADJ MEANS	1	193.43			
EQUAL SLOPES	4	552.99	138.25	.92485	.4488
ERROR(EACH REGR)	723	.10807 +6	149.48		
- - - - - - - - -					
TOTAL	732	.13099 +6			

TABLE OF COEFFICIENTS

COVARIATE	COEFF	STD ERROR	T-STAT	SIGNIF
11.CHLT15	-1.3163	.23838	-5.5218	.0000
406.AGE	-.88667	.86335 -1	-10.270	.0000
706.AGESQ	.11254 -1	.11732 -2	9.5924	.0000
839.INC.SUB	.39561 -2	.10738 -2	3.6842	.0002

TABLE OF MEANS AND REGRESSIONS

PROGPART	YES	NO
MEAN	93.520	94.335
ADJ MEAN	94.335	93.223
(STD ERROR)	.61212	.72246
INTERCEPT	109.75	107.64
N	423	310
CONSTANT	107.14	109.62
11.CHLT15	-1.2617	-1.3927
406.AGE	-.75020	-1.0699
706.AGESQ	.94105 -2	.14108 -1
839.INC.SUB	.35087 -2	.40460 -2
SE OF REGR	11.419	13.253
R-SQR	.13154	.21353
SIGNIF	.0000	.0000

dependent variable above and beyond that explained by the covariates. Here we accept the null hypothesis of no effect--participation is not significantly related to nutritional status. The third and fourth tests determine the nature of the effect (non-existent in this case) of the qualitative variable on the relationship between the dependent variable and the covariates. In short, these two tests look for changes in the regression between the dependent variable and the covariate for the different groups established by the qualitative variable. The first of these tests looks for shifts in both the coefficients and the intercepts of the separate regressions; the second checks for shifts in the coefficients alone. In our example, participation does not significantly cause any such shifts. (Shifts in the coefficients can be thought of as changes in the regressions due to the interaction of the qualitative variable and the covariates.)

General Comments on Parametric Tests. The statistical techniques described above are all "parametric"; they assume that some of the variables in the analysis take on known distributions which can be specified by knowing one or two parameters of the distribution (for example, a normally distributed variable is specified if its mean and variance are known). Additional tests are available to verify the validity of the assumptions, and procedures have been developed in some cases to account for violations of the assumptions in the analysis. Usually, when an assumption is violated, the compensating procedure reduces the power of the test (making it more difficult to reject the null hypothesis), and always makes the testing procedure more complicated.

In the next section we will discuss some nonparametric alternatives to our parametric tests. Note that our examples of analysis of variance and covariance were both "one-way" models; only one qualitative variable was used. Somewhat more complicated two-way models exist and can be used when appropriate.

Non-Parametric Tests

Non-parametric tests can be applied in two situations where parametric tests fail: (1) where the dependent variable is not quantitative; and (2) where the assumptions of the parametric test are not met. They are essentially distribution-free and often use rankings within a variable rather than numeric differences between "scores."

Chi-square Test for the Existence of a Relationship. The simplest non-parametric test is the Chi-square test of the existence of a relationship. If the data from two variables fall in discrete categories, this test can be used to determine whether a relationship exists between them.

Consider the discrete classification of nutritional status often used in nutrition programs and studies. (This is a partition of the quantitative percent of standard variable used earlier.) Suppose we wanted to know if participants differed from nonparticipants with respect to the relative frequency of occurrence of the different nutritional grades in each group. Table 5-16 is the two-way or contingency table presenting this data for Candelaria Revisited.

One of the MIDAS outputs below the table is the value of Chi-square and its significance level. Selecting either of the classically accepted significance levels of .05 or .01, we can reject the null hypothesis of no relationship; that is, nutritional grade is related to program participation.

Note that our t-test for difference of means suggested that the means of the quantitative representation of nutritional status (percent of standard) were not statistically different. The Chi-square test suggests that overall, the distribution of nutritional status in participants is statistically different from the distribution for nonparticipants. This should serve as a friendly reminder of several of the points made in our analysis section. The selection of variable (quantitative or qualitative) affects results. Also the selection of test (means test, overall distribution test) affects results.

TABLE 5-16

A SAMPLE CHI-SQUARE TEST;
CANDELARIA REVISITED

TWOWAY CROSS-TABULATION

500. PROGPART		445.NUT.STAT MISS	2AND3	FIRST	NORMAL
N=	848				
TOTAL=	859	11	55	169	624
YES	487	5	22	108	357
EXPECT			32	97	358
NO	361	6	33	61	267
EXPECT			23	72	266

TESTS OF INDEPENDENCE	STATISTIC	SIGNIF	DF= 2	N= 848
MAXIMUM LIKELIHOOD	9.6948	.0078	CRAMER'S PHI=	.1
CHI-SQUARE	9.7452	.0077	CONTINGENCY COEFF=	.1

Mann-Whitney U Test. Of course, we should realize that the Chi-square test is not testing quite the same thing as the t-test for difference of means. A closer parallel exists between the t-test and the Mann-Whitney U test. This test is a "ranking" test, and the variables must be at least measurable on an ordinal scale.

Table 5-17 gives the Mann-Whitney U test for our Candelaria Revisited data, using the percent of standard measure of nutritional status. The test compares the ranks actually attained by one group (nonparticipants) in the entire data set with the ranks one would expect to find if the groups were equal. If the ranks differ, the groups are not the same. Our result for Candelaria Revisited is that the groups—participants and nonparticipants—are not statistically different.

Kruskal-Wallis Test. Multivariate non-parametric analogs to the parametric techniques also exist. For example, the Kruskal-Wallis test performs a one-way analysis of variance using a ranking technique. This particular test is used to determine if "k" independent samples are from different populations. Table 5-18 is the MIDAS Printout of the Kruskal-Wallis test for Candelaria Revisited data using the revised definition of participation distinguishing between children under two and children over two. The result parallels the one-way analysis of variance. The groups are in fact different as indicated by the .0000 significance level on the printout.

Review Results with Project Personnel

During the course of our project, we were struck repeatedly with the importance of including project personnel in the interpretation of analysis results. After analyzing data from Honduras, it became obvious from comparing characteristic curves of new program participants with continuing program participants that new participants entering the program in or after the ninth month were "better off" with regard to nutritional status than their continuing counterparts. We formed competitive explanations for this observation; for example, the method of recruitment

TABLE 5-17

A SAMPLE MANN-WHITNEY U TEST;
CANDELARIA REVISITED

TWO-SAMPLE COMPARISON

TEST OF 443.PCW76		SIGNIF
MANN-WHITNEY U=	84639.	.3545
MEDIAN TEST		.3302

PROGPART	N	AVG. RANK	MEDIAN= 93.000 N<	N>	N=
YES	487	417.797	242	229	17
NO	361	433.543	173	182	6
TOTAL	848 OUT OF 859				

TABLE 5-18

A SAMPLE KRUSKAL-WALLIS TEST;
CANDELARIA REVISITED

MULTI-SAMPLE COMPARISON

TEST OF 443.PCW76	STATISTIC	DF	SIGNIF
KRUSKAL-WALLIS	23.034	3	.0000
MEDIAN	15.387	3	.0015

PROGPART	N	AVG. RANK	MEDIAN= 93.000		
			N<	N>	N=
GT2CP	308	420.508	148	147	13
GT2NP	276	378.817	159	113	4
LT2FP	113	457.345	48	61	4
LT2NP	151	491.563	60	89	2

TOTAL 848 OUT OF 859

of new participants found all those children who were in worst shape first, or, alternatively, as program participants were drawn from a wider geographical area, a healthier set of participants was found. The "truth" was learned only after showing our analysis to the project staff. They knew--prior to our analysis--that we would observe this phenomenon; the government doctor in the region ordered them to concentrate on grade II malnourished children because the most malnourished (the grade III children) were beyond help.

Similar examples can be drawn from KOTTAR, Esperanca, and Indonesia. In every case, the intimate knowledge of the program staff with their community was essential for interpreting analytic results. In KOTTAR, the staff recognized the effects of the drought on the data. In Esperanca, the staff recognized the mistaken inclusion of children with birth defects in the data set. In Indonesia, the staff disputed our explanation that the environment deteriorated and forced us to look elsewhere to explain an apparent lack of program impact. (They also provided us with an acceptable alternative: the program was too new and had been interrupted by government military action directed at the staff itself.)

Iterate Analysis and Review Steps

Except in extraordinary cases, a single pass through steps 3 and 4 will not be sufficient to limit the competing explanations of observed outcomes. Our discussion of sophisticated analytic techniques indicates how much judgment and art is involved in social science analysis. A single attempt to perform and interpret analysis hardly allows for the creativity of the analyst to lead to truth.

Incorporate Decisions into Local Decision Processes

Although there are certainly situations where learning for learning's sake is laudable, we believe that learning for program improvement is best. Efforts should be made by both the analyst and program leaders to modify the intervention, the data collection, and the data analysis procedures in response to whatever is learned through analysis.

CHAPTER VI
ELEMENTS OF INTERVENTIONS THAT CONTRIBUTE TO SUCCESS

Thus far, this report has dealt exclusively with methodological issues: problems of gathering, analyzing, and interpreting data. We now attempt to apply the findings concerning methodology to the data used during this project. This chapter addresses directly the first stated objective of this project: identifying the characteristics of community-based nutrition interventions that contribute to their success.

There are three types of conclusions regarding characteristics of interventions that enhance success. First, there are general conclusions applicable to any social intervention, but of particular importance with regard to nutrition programs. Next, there are some more specific conclusions concerning the "process" of nutrition interventions. Some of these follow directly from the general conclusions and are discussed specifically in the context of nutrition. Others are peculiar to attempts to intervene in the developing world, and are also placed in the context of nutrition. Finally, there are comments (as we shall see, conclusions may not be the correct word) concerning the specific components often found in nutrition interventions.

GENERAL CONCLUSIONS

> The setting in which community-based interventions are carried out will most likely change in major ways during the course of those interventions.

This point has been raised repeatedly throughout this report. This is not an accidental redundancy; we have been struck repeatedly by the importance of the changing environment. Therefore, once again, examples of this phenomenon are included.

From the literature, the most dramatic example of a changed setting comes from Urrutia, who describes the unexpected termination of his corn-fortification intervention after an earthquake killed 5 percent of the study population and destroyed nearly all the homes in the study region (Urrutia, 1976). That type of drastic change can end interventions. Yet more subtle changes in the environment occur also which can be accommodated during an intervention.

In Thailand, Dr. Gershoff described the near collapse of the rice market near Chiang Mai due to the black market movement of rice across the country's borders. In Esperanca, manpower availability is constantly changing as the number of students from the University of Para working with the project in Santarem fluctuates. In KOTTAR, the monsoons failed to appear in 1975, which curtailed production and resulted in the doubling of the local price of rice and tapioca (cassava). The Save The Children development programs in the Dominican Republic and Honduras have had to adjust to changes in leadership, as SAVE transfers its personnel among programs. Recently, PRIMOPS has changed its mix of services and expanded the geographical area served in response to its need to secure additional funding.

The point of these examples is that shifts in circumstances are neither minor nor insignificant. Furthermore, to the extent that community-based interventions focus upon high-risk communities or subgroups within communities, exogenous events, particularly floods, food shortage, droughts, etc., generally affect these high-risk groups the most. Thus, even moderate changes in circumstances can play a significant role on variables used to measure outcomes while major changes may dominate them.

The second general conclusion reflects the general level of knowledge held by the nutrition/health planning field.

> While the state of the art of community-level problem diagnosis has improved, it offers at best **only a first approximation** to an understanding of the specific causes of malnourishment at the local level.

In recent years considerable attention has been given to trying to

understand the specific physiological determinants of malnourishment. These include:

- the role of gastro-intestinal diseases in malabsorption (e.g., stealorrhea, sprue, Crohn's Disease, ulcerative colitis),
- infective diseases (e.g., respiratory, tuberculosis, leprosy, yaws, scabies, sepsis, cardiace),
- allergies (e.g., food asthma, eczema, celiac disease),
- endocrine and metabolic diseases (e.g., diabetes, Addison's Disease, neoplasms),
- congenital defects (e.g., prematurity, metabolic errors, anatomical, nutritional),
- various results of contaminated foods, including stomatitis, enteritis, dysentery, parasites, and toxins (Robson, 1972).

Considerable progress has been made, especially at the clinical level, in understanding these determinants. For example, Scrimshaw et al. outline the interaction of infectious disease and malnutrition (Scrimshaw, Taylor, and Gordon, 1968). In-depth examples of the synergistic and antagonistic interaction at the clinical level are provided. Bacterial, viral, rickettsial, and other types of infections are discussed in reference to various nutrient deficiencies. More recent works, one edited by Suskind (1977), and the other authored by Chandra and Newberne (1977), also address the infection-malnutrition interaction. They substantiate the great strides made in understanding the determinants of malnutrition at the clinical level.

Numerous other documents and research findings concerning physiological determinants of malnutrition for at-risk groups, such as women and children, also exist. For example, Fomon (1977) describes how to identify disorders with a nutritional basis in children, and Jelliffe (1975) concentrates on issues of infant malnutrition in developing countries. Similarly, Moghissi and Evans (1977) present papers of scientists, clinicians, and investigators who are exploring the nutritional needs of women and their dynamics to assure optimal reproductive performance, and a publication edited by Mosley (1978) deals with relationships between

maternal nutrition, reproduction, and the health and welfare of their offspring. These, and numerous other publications, are indicative of the sophisticated understanding of the physiologic dynamics that determine malnutrition and must be addressed.

In addition to advances at the clinical level, much work has been done to contribute to the understanding of the role played by environmental, social, and economic factors in causing malnutrition. In the environmental area, the importance of sanitary facilities in reducing diarrhea has been shown. Rubenstein et al. showed significant reduction in hospital admissions for diarrhea in a Hopi pueblo after the construction of sanitary facilities (Rubenstein, et al., 1969). With regard to social factors, Jelliffe notes how food selection and availability, culinary practices, and other food-related behavior are transferred to individuals through a socialization process (Jelliffe, 1975). Jelliffe argues that an understanding of the cultural factors in a society can lead to more directed efforts at combatting malnutrition. Pollitt, Gilmore, and Valcarcel studied the behavior of mothers and infants during feeding and concluded that food intake of newborns might be improved by "teaching" mothers better behavior patterns while nursing (Pollitt, Gilmore, and Valcarcel, 1975). With regard to economic factors, many studies have been done in an effort to relate income, occupation, wealth, etc., to nutritional status (Drake and Fajardo, 1976; Heller and Drake, 1979). Ifekwunigwe brings land tenure, production methods, and income to this analysis (Ifekwunigwe, 1978). Lechtig created a socioeconomic scale based on housing, clothing, and child education, which had an inverse relationship with birth weight (Lechtig, 1975). Also, Valverde related amount of land owned and occupation to the prevalence of malnutrition in children (Valverde, 1977).

As good as some of these analyses are, an intervention based on knowledge derived from them can, at best, provide a first approximation to the best strategy for a given situation. Any given local system is sufficiently complex that a model or theory of local change based on existing knowledge will be incomplete. Based on analysis of local data, it might be possible to conclude that due to a high initial incidence of

diarrheal disease, resource allocation to improve water potability might be a wiser start than allocation to food supplementation. But the causes of diarrhea may be multiple, and water potability may be only an initial step toward eliminating the disease. Typically most interventions have too few resources at the outset to perform an analysis to arrive at even this first approximation.

It is easy to become discouraged by a discourse which emphasizes what we do not know and what we cannot do. But, a first approximation is fundamentally better than none at all, if it is understood that it is just a first approximation and not an ultimate solution. Certainly, as complete a diagnostic process as possible ought to be carried out in any given setting. The principal point is that this diagnosis is, and will continue to be, preliminary and tentative.

The third general conclusion is a direct consequence of the first two.

> It is virtually impossible to plan **in advance** what will be the optimal intervention strategy. Rather, an initial plan incorporating mechanisms for modifying the intervention as time progresses will produce improved courses of action for achieving the ultimate goal.

The logic underlying this conclusion is relatively straightforward. If the setting containing an intervention is constantly changing and if the state of knowledge is such that only a first approximation to a solution can be found at any point in time, a strategy of iterative programming is needed to enable proper response to shifts in setting and/or to lessons learned while trying that first (and subsequent) approximations. In short, the world is complex--we must be flexible in order to deal with the complexity. This may sound like the typical platitude of the social scientist who must explain his/her inability to effectuate positive change. But it is true and, if it is recognized and dealt with intelligently, its tendency to confound success can be minimized.

We are particularly concerned with the unanticipated responses of a complex society to externally supplied stimuli, such as interventions. Because most of the basic determinants of malnutrition are intricately

bound to the whole fabric of a community, the responses to attempts to modify those determinants have far-reaching consequences.

As illustration, consider some of the now recognized but once unanticipated consequences of some of the more intuitively appealing solutions for malnutrition. Supplementary foods can have the effect of depressing prices paid to farmers which, in turn, causes decreased local production. Isenman and Singer found less of this disincentive in India than they originally thought they would, but they did find evidence of it (Isenman and Singer, 1977). Latrines, when built in the wrong circumstances (in areas without fly control or where water tables are high) can do more to increase the spread of infectious disease than they do to decrease it. As a final example, the mere presence of an intervention can reduce the willingness of local communities to help themselves by creating an unnatural dependence upon outside aid. A CSF team reviewing PL 480 programs in India writes:

> The assessment team was struck time and again by the recipients' unquestioning assumption that the government would provide food, schools, wells, and shelter. We believe that this dependence is furthered when food aid is not made conditional on the availability of local inputs (Community Systems Foundation, 1979).

We also note that as some of these unanticipated relationships become "anticipated," and as other better known relationships become documented, they take on complex behaviors of their own. As consumption increases, malnourishment decreases (usually). But this relationship is not yet sufficiently quantifiable. "Magnitudes" of improvement in nutritional status cannot be associated with corresponding "magnitudes" of change in consumption. There are diminishing returns for each increment of consumption and, at some point, further increases become a liability in the form of overnourishment leading to obesity.

In summary, planners of nutrition interventions are faced with: (1) bivariate relationships which are themselves complicated and nonlinear; and (2) multivariate interactions which are complicated and often unanticipated. The existence of complexity and uncertainty renders the construction of a

single "model" or "mathematical representation" of reality a difficult if not impossible task.

Such a model would need to consist of a set of simultaneous, nonlinear equations with interaction terms. Most of these relationships would also have substantial variablity associated with them, indicating that a stochastic treatment would be most appropriate. Models of this type are mathematically unsolvable. Thus, community reality cannot be abstracted even if data are available, and they are not. Under these conditions, it is clearly difficult to know in advance the best way to reduce child malnourishment at the community level. Therefore, the most reasonable strategy is to formulate an initial plan and embark upon a course of community action, while incorporating mechanisms for modifying the intervention as these intricate relationships become understood. (An initial attempt to capture nonlinearity, interaction, and simultaneity in some of the relationships is reported in a model shown in Heller and Drake (1979) and Drake and Fajardo (1976). Several heroic simplifing assumptions were needed to structure a model amenable to manipulation.)

THE PROCESS OF A SUCCESSFUL NUTRITION INTERVENTION

The general conclusions of the last section call for a fluid and flexible approach to planning and implementation. Programs that recognize this fact and design their approach accordingly have a higher likelihood of success in reducing malnourishment than others. This section describes five characteristics of an intervention that contribute to its ability to identify and respond to a changing environment.

Articulation of a Theory of the Local Nutrition System

> A nutrition program that articulates a theory of the "local nutrition system" will be better prepared to identify critical relationships than one that does not.

The "local nutrition system" is the collection of primary and secondary factors which relate to the nutritional status of a community. For example, CSF has used the nutrient flow model as an initial theoretical

description of the "nutrition system." This model recognizes consumption and health as two primary factors in any nutrition system, and includes income, agricultural production, and sanitary systems as secondary factors determining consumption and health. This general framework is a useful starting point for articulating a "local nutrition system" theory. The specific concerns of the local situation must be incorporated into the general framework.

For example, the Cauca Valley in Colombia (the site for Candelaria and PRIMOPS) is virtually seasonless. The slight variations in rainfall which distinguish the dry and wet seasons are not sufficiently significant to bring "seasonality" into the local model. In Thailand and India the monsoon season is definitely a factor to be reckoned with; it must be in the local model. Similarly, in the Central Amazon, there are local taboos against eating certain nutritious foods during pregnancy. A different set of taboos exist in Indonesia, while in India, some women choose to cut down on consumption during the last three months of pregnancy because they believe that it is easier to deliver small babies. These local customs must be made a part of the articulation of the "local nutrition system."

It should be immediately apparent that the usefulness of the local model lies in its ability to help test "a priori" the impact of a particular intervention strategy. However, there are two other uses for the model. First, it serves as a guide to the data to be collected for monitoring, and an aid in setting reasonable expectations of the impact of proposed strategies. (A theory acknowledging the sharing of food within a family will lead a change agent to expecting less change in nutritional status of children as a result of supplementation than a theory which ignores this phenomenon.)

Learning from Experience

> A nutrition program capable of monitoring and evaluating itself is more likely to reduce child malnourishment than one that does not have these characteristics.

This statement sounds trivial, but our experience indicates clearly that

the implementation of this concept is not. In the interventions considered in this project, the objective data gathered during the intervention was generally underutilized by the primary change agents for self-evaluation.

In order for data to be used in monitoring and evaluation, there must be a clear recognition of the usefulness of such feedback by project personnel. In every project studied, this interest and recognition was present to some degree. Along with it, however, was the perception that the gathering of data on project performance is a diversion of resources away from the provision of services to beneficiaries. The question becomes one of cost versus benefit and, here, the "show me first" philosophy often wins out.

The interventions studied as part of this project were sufficiently committed to take the first step; to routinely record anthropometric and, in some cases, additional data as part of their everyday activity. However, in several cases, this commitment to data gathering was not matched with sufficient analysis and feedback to project staff. We believe that this is the result of several forces. First, the costs of analysis and "feeding-back" are grossly underestimated. Second, the skills required to perform analysis and provide feedback are often absent from the staff of a community-level program.

With regard to the costs, data collection should be only a small part of the total effort. Detailing percentages to be spent for data gathering, cleaning, analysis, and feedback in this report is a risky undertaking. Still, we offer some guidelines. Except for interventions launched for strictly research purposes, it is difficult to imagine conditions which justify the expenditure of more than fifteen percent of total project resources on monitoring and evaluation. A more reasonable estimate is closer to five or ten percent. Of this amount, 20% should go toward collection, 20% toward cleaning, 30% toward analysis, and 30% toward feedback. We point out that a $100,000 per year intervention would spend only $3,000 on data analysis.

The 30% figure for feedback might seem on the high side, but we have found that the most difficult aspect of feedback is presenting results in a

useful manner. (The average change agent has trouble interpreting an analysis of covariance—so, by the way, has the average analyst.) Great care must be taken to feed back data correctly, for it is difficult to achieve proper balance in the use of quantitative feedback. Often, the first time that numeric results are provided to a community, there is a tendency to over-subscribe to their importance. In Honduras, there was consideration given to stopping a feeding program because a brief field study showed a slight increase in malnourishment--statistically **insignificant**, but a change for the worse nonetheless. A drought was worsening, but this was not immediately apparent in the field. This bias in favor of numeric results is especially distressing because so many of the most important phenomena are not presently capable of quantification. Careful presentation of results is the only hedge against such misinterpretation.

With regard to analytic skills, these might best be supplied by an outside agent. This is especially attractive if that outside agent could provide service to multiple communities, for his own skills would sharpen, and economies of scale would be derived. If, as we shall recommend, the locally generated data sets are used as part of a national surveillance system, one function of the national team running the system could and should be the analysis and feedback of results to participating communities.

Local Involvement

> Nutrition interventions that emphasize local involvement, both in design and implementation, are more likely to succeed than ones that do not.

The literature is replete with examples where lack of local involvement is faulted and the presence of it praised, yet differences of opinion exist with regard to the degree and nature of involvement. Also, the definition of "local" can be a source of contention. The following pages argue in favor of a dominance of local input in both the design and execution of community-based nutrition programs.

We begin with a definition of the term local. A local person lives in

the geographically defined community and would continue to do so in the absence of the intervention. An argument can be made that government officials and/or medical people should be considered a part of the community for which they have moral and professional responsibility. We choose to regard these people as an essentially external resource; that is, they can play a special role as "pseudo" community members who provide needed assistance, but they do not represent the community itself. Given this definition, only one intervention examined in the project had a strictly local change agent. In KOTTAR, the Belgian priest and sister had lived in the region for a long time prior to the intervention and planned to die there. In Candelaria and PRIMOPS, the change agents were predominantly external: medical people with official responsibility for the study area but who did not live there.

There are three major reasons for encouraging local involvement. First, outside change agents generally know too little about the environment in which they are working. Though they are often knowledgeable about development, medicine and/or nutrition, they lack in-depth knowledge of local custom and culture. Local involvement brings with it the intimate knowledge of local tradition that must be factored into any intervention plan.

Second, it is difficult to grasp the complex political structure and often transparent hierarchy at work in a community. Initial identification of key persons must often be altered, sometimes only after an intervention has gone astray. The answer, again, is a high degree of local involvement.

Finally, a key to success is acceptance. The best plan will not work if the beneficiaries are unwilling to participate. Local involvement in planning, implementation, and evaluation is an important step toward gaining this acceptance.

A logical question is, how can local expertise be combined with the knowledge and funding arriving with the outside change agent? We have observed several different styles and conclude that there is no single correct way. The personality, skills, and station of the change agent dictate his/her approach; what works for a minister may not work for a

doctor. Save The Children, for example, has a conscious and deliberate strategy for enlisting commuity involvement—CBIRD, or Community Based Integrated Rural Development.

The first step in this strategy is the establishement of a broad-based community development committee which works with the SAVE project leader to define a workable program. In practice, incentives for undertaking certain types of activities may be tied to the skills of the SAVE staff and/or the availability of outside funds. The clear advantage to this approach is that the intervention becomes a community program, and is likely to be accepted and continued. The disadvantage is that participation at all stages of the intervention is time-consuming, and programs assembled by committee tend to have patchwork qualities which, although representative, impede efficiency.

In contrast, there is the approach used by the Aroles in Jamkhed, India or, to some extent in India by the KSSS. In both locations, there are too many villages for the available resources. The change agents select those villages that are most eager to participate and most willing to accept the solutions and/or "rules" set by the change agent.

Whatever the method, local involvement is generally beneficial. The one caveat with regard to participation is that the community is often represented by an existing power structure with a stake in preserving itself and the gap between it and the poorest classes. Care must be taken to avoid the situation where the intervention strengthens the hold of the powers that be, at the expense of the more high-risk people in the community.

The Bridge Between Two Worlds

> A nutrition intervention director who spends as much time as possible in the community, yet attempts to promote linkages with non-local systems, raises the likelihood of reducing malnourishment at the community level.

In the discussion of the life cycle of an intervention, we spoke of the change agent as the bridge between two worlds—the community and the

world of international government and finance. We made reference to the role the change agent has in bringing the needs of a changing community into harmony with the technical knowledge and resources of the outside world. The double-edged premise above suggests that capable performance of the bridging function is a key to attaining the ultimate objective.

Let us consider linkages to the outside world. At the very least, the program director must secure funding during the initial (and, perhaps, later) stages of his intervention. However, it is the change agent who brings the technical knowledge and organizational skills to an intervention. He can also bring potentially available government services to a community and can serve as a catalyst in seeing that existing services are dispensed effectively.

The models for achieving effective linkages vary according to program design, individual personality, and organizational support. Save The Children (the advocate of community development committees), often locates the primary program director (person with the power to make policy) outside of the study area in a national or regional city. Operational responsibility of the program is usually vested in a native born person who resides in the study area itself. In KOTTAR, the KSSS connection with the Catholic Church and Catholic Relief Services forms a highly successful link with a source of the external support, while the primary change agents remain in the study area most if not all of the time. Candelaria and PRIMOPS are allied to a major university, which provides technical talent as well as a financial bridge to the outside.

Now we turn to the change agent's presence in the community. In more remote areas, maintaining an adequate presence in the community while establishing linkages with the outside world are often conflicting activities. Too often, for both personal reasons and "to keep the money flowing," the program director feels compelled to locate in a major city. Programs can stagnate in these circumstances unless extraordinary efforts are taken.

Again, there is no magical formula for maintaining a balance between the outside world and the community. It is a matter to be resolved in

each case, yet success of the intervention may well hinge upon resolution.

Ensuring a Critical Level of Intensity

> A multi-faceted nutrition intervention offers a variety of services to its beneficiaries. Each service must be maintained at some critical level sufficient to produce the desired result.

Of the five characteristics of a successful nutrition intervention, this one speaks most directly to the activities or services that comprise the intervention. Although it is intuitively appealing to speak of a critical level of intensity (a threshold of activity needed to produce a result), it is difficult if not impossible to quantify that critical level for a given activity. Therefore, we must rely upon intuition to provide insight.

First, consider a home visit program. In KOTTAR, visits were monthly; in Candelaria, visits were made every second month; in PRIMOPS, visits were six to eight months apart. Clearly, not all rates of visitation produce the same response. In assessing optimal frequency of visitation, the quality of the visit, the skill of the home visitor, and the nature of the local malnutrition problem itself must be considered. We feel that home visits for a program using weight/weight(age) measures as a diagnostic and/or yardstick of change should occur every three months. With the exception of the highest-risk children who dehydrate quickly (in hours sometimes), three months is enough time to see change in the measure.

Second, consider a supplementary feeding program. There must be a critical amount of food to produce a visible change in nutritional status. Again, this critical amount will vary from place to place, depending on the magnitude of food shortage, the amount of intra-family sharing, and the disease rates in a communtiy. In this case, we will not even speculate as to the critical level; it is sufficient for our purposes to state that it exists.

As a third example, consider the problem of identifying critical levels in a multi-dimensional program. The interaction between the dimensions affects the critical levels of each. Diagnostic home visits, which also

deliver food supplements, must be frequent in order to avoid spoilage of the supplement and to facilitate movement of the food.

In summary, we have identified five characteristics of a successful nutrition intervention which facilitate the incorporation of a flexible process of program planning and implementation. A nutrition intervention director should formulate a theory of the local nutrition system, develop the capability to learn from experience, encourage local involvement, maintain a distinct presence in the community, yet establish linkages to the outside world, and strive to provide services at their critical level of intensity.

Yet there is a major reason why many programs funded from outside sources, such as US/AID, do not subscribe to this philosophy. A process-oriented model of intervention is inconsistent with the budgetary cycle of governmental, quasi-governmental, and large voluntary organizations. Funders are obliged to commit in advance specific amounts of resources for distinct purposes, and are themselves accountable to some higher review. Thus, they find it difficult to accommodate the uncertain pace and substance of interventions which adopt flexible procedures to enhance internal modification, rather than chart in advance a specific regimen.

There is, however, an additional irony regarding the process perspective of an intervention, and it causes particular concern. If a distinct plan is laid out in advance and adhered to during the funding period, it becomes relatively easy for an outside analyst to evaluate the project. Clearly, a very real need exists for funders to evaluate interventions; yet, at first glance, the process model makes this a potentially more difficult task.

However, we believe this problem dissipates when approached through the iterative process of participatory evaluation. An iterative approach enables the evaluator to become acutely aware of changes in the local environment as well as intimately involved with project personnel, and hence develop an understanding by which to assess the viability of the intervention. The same process which allows project directors to effectively guide their program can give outside reviewers the means of assessing their "investment."

SUBSTANTIVE INTERVENTION CHARACTERISTICS

In this chapter, we have argued that the components of a given intervention should be uniquely determined through the application of a flexible "learning" process of planning, implementation, and modification. This is not to imply that an intervenor cannot or should not learn from the experience of others. In this section, we summarize our interpretation of the experiences of the change agents with whom we have worked throughout this project.

In the introduction to this report, we presented a broad definition of a nutrition intervention, one that recognized a range of possible activities including, on one extreme, the intuitive solution of supplemental feeding and, on the other, the less obvious solution of general development work. In fact, the interventions reviewed in this study almost always used a combination of components applied at varying levels of intensity. Attacks on malnutrition for immediate relief were often combined with attacks on poverty for long-term, self-sustaining relief. Table 6-1 categorizes the components of each of the interventions for which we have anthropometric data.

For convenience, our discussion describes five broad categories of activities. However, note that these often overlap in practice. Before we begin, it is important to lay one issue to rest: the distinction between curative activities and preventive activities. There is no clear demarcation between the two; what is curative in one context is preventive in another. The rehydration of diarrhea victims is curative with respect to diarrhea and preventive with respect to malnutrition. The feeding of pregnant mothers is curative for the mothers but preventive for the infant (by leading to higher birth weight). Because of this high degree of overlap and confusion, we choose to obscure the distinction between prevention and cure—yet at the same time emphasize preventive aspects of interventions.

The following section describes our current understanding of the substantive components of interventions.

TABLE 6-1

CLASSIFICATION OF INTERVENTION COMPONENTS BY PROGRAM

Name of Program	Improvements in Food Consumption								Improvements in Health								Improvements in Environ. Conditions					Family Planning	
					Dist.				Prevent. Health			Curative Health		Weight Charts					Waste Disposal			Family Planning	
	Fortification	Supplementation	Production	Storage	Among Families	Within Families (educ.)	Nutrient Education	Increasing Disposable Income	Immunization	Hygiene Education	Rehydration Program	General	Malnourishment Recoup	Kept by Mother	Kept by Local Ctr.	Kept Centrally	Improved Potability	Improved Accessibility and Quantity	Latrine Constr.	Education	Sewage/Drainage	Suppl. Education	Prov. Contraception
Esperanca (4 villages)							+		(+)	+	+	+	+		+		(+)	(+)	(+)	+			
Candelaria						+	+		+	+	+	+	+	+	+		(+)			+			
Candelaria revisited						+	+		+	+	+	+	+	+	+		(+)			+			
PRIMOPS (4 regions)							+		+	+		+	+		+		(+)			+			
Honduras (5 villages)		+	+			+	+	+							+			+	+	+			
Tangse (11 villages)			+	+		+	+					+			+				+				
Thailand (29 villages)	+					+	+		+	+		(+)				+							
Kottar (21 villages)		+	+		+		+	+								+			+			+	
Dominican Republic		+	+			+	+	+			(+)	(+)							(+)				
CARE		+												+									
Papago Indians		+					+		+	+	+	+	+			+			+	+		+	+

KEY: + = Components implemented by intervention programs

(+) = Components implemented by other organizations

Food Consumption

All of the interventions studied during this project included some method for increasing food consumption. These ranged from education directed toward improved use of available foods, to production of new foods, to outright supplementation and/or fortification. Our basic position is that adequate food intake is a necessary but not sufficient condition for good nutrition. Furthermore, efforts to increase consumption through supplementation programs should never be thought of as a long-term solution to the problem of malnutrition. At best, supplementation can temporarily alleviate a critical shortage and, even in this situation, improved consumption is still not sufficient for good nutrition. In our study, we were unable to assign direct positive effects to any of the other strategies used to bolster consumption.

An unrealistic view of the prospects for food supplementation is frequently held. In practice, rations are too small to have much impact, especially since they are almost always shared among family members. We found sharing of the supplement even in on-site feeding programs; for example, in the India ICDS and Jamkhed programs, food in the form of buns or plate servings was taken home from the site for sharing with siblings.

Unfortunately, the factors that might "prove" that supplements cannot work are themselves unknown. The U.N. University, WHO, and FAO are currently engaged in a large study to redetermine daily nutrition requirements. Furthermore, the effects of disease upon nutrition requirements are unknown. But, our judgment at this time is that supplements alone do not substantially affect malnutrition.

We do acknowledge that supplements can be used for another important purpose, namely, as an incentive for participation in nutrition-related programs. KOTTAR used food obtained through Food-For-Work as an incentive for development projects and Food-For-Peace commodities to induce family acceptance of their home visits. In Honduras, the size of the program grew at a rate far greater than the rate of growth of the supplement, simply because people wanted free food. In the context of

incentive to participation, food supplementation is helpful.

Health

With the exception of the CARE/PL 480 food supplement, every community-level intervention studied in our project had a health improvement component either as part of that intervention or as part of another related simultaneous intervention. The precise nature of the health component is very much a function of the local environment, of the medical training of the intervention staff, and of the medical treatment facilities and people already in a community that can be tapped by the intervenors. Therefore, it makes little sense to prescribe a "best" health component. However, there are four activities that are certain to have positive but as yet difficult-to-quantify effects upon participants.

The first activity is the reduction of diarrhea. The relationship between malnutrition and diarrhea is well-documented as well as intuitively appealing. (A child whose digestive tract is not functioning needs more nutrient intake.) The most common solutions for diarrhea problems at the community level, such as the construction of sanitary facilities and improvements in water potability, represent environmental improvements and are discussed in the next section. The most common solutions at the individual level are hygiene education and oral rehydration. This latter solution, the treatment of diarrhea by oral rehydration, is simple and effective, and also requires hygiene education. Although solutions to the diarrhea problem span the five categories we chose to discuss, the problem is sufficiently important to rate special attention as a health measure.

The second activity is immunizations. The third is the administering of broader curative services for diseases other than diarrhea. These two activities are addressing the same objective—to minimize diseases that have synergistic relationships with malnutrition.

The fourth activity is the use of a growth chart for participating preschoolers. It is the best diagnostic device we have—imperfect as it is. Because growth charts are intricately involved with the education of mothers, this will be discussed further in category 5.

Environmental Works

We have already alluded to the construction of sanitation and water systems as a strategy for minimizing the spread of disease, especially diarrhea. Yet there are important qualifications on any recommendation to build these improvements.

The construction of latrines is the most common response to widespread diarrhea. However, in some situations, latrines can do more harm than good. Latrines may do more to spread disease than prevent it if: (1) there is not fly control around the latrines; (2) the water filters from the latrine into the water supply; and (3) people do not learn how to use them properly. In Jamkhed, latrines were not utilized because of the fly problem. In Indonesia, the government built latrines but people were never instructed in their use, and they were not used.

A typical response to non-potable water is to teach people to boil it, or to alter their source. It is not clear that either of these solutions is best. There is some evidence that the quantity of water used is more important than water quality itself; flushing the home environment by using more water, even if it is of poor quaity, may be more important than cleaning that which is used.

Family Planning

Most of the interventions studied did not have a family planning component. However, in Candelaria, we did obtain some rather startling results which merit special attention, not only because of the family planning implications but also because of what it may say about intervention components in general.

Throughout the entire Candelaria nutrition project (1968 to present), there was a government program in family planning occurring in the city of Candelaria. The program consisted of distributing educational materials, as well as providing contraception consultation and prescription. This local component of a national program seemed to be well run and funded adequately for the tasks it was intended to carry out. The Promotora

Program enhanced this national program by providing an "outreach" component by disseminating material and following up problems during the course of the promotora visits to the home. During the course of the intervention the use of birth control increased dramatically, but it was not possible at that time to discern the effect upon the fertility rate of mothers. However, the Candelaria Revisited effort provided us with an opportunity to look at the longer-term effect of contraception use upon fertility. In the study, it was found that those mothers who participated in the Promotora Program had much lower fertility rates than those who did not, even though both groups had been exposed to the same government program. Clearly, a more intense application of the contraception program, in the form of promotora outreach and follow-up, yielded dramatic returns beyond the effect of the government program alone.

Nutrition and Hygiene Education

The importance of nutrition and hygiene education can be neither ignored nor quantified at present. We have already alluded to its role in changing consumption patterns, in treating diarrhea through oral rehydration, and in teaching people to use latrines and contraception. Although we cannot place a "number" on the impact of education, we feel that its role as a catalyst may be the most critical component of any intervention.

Yet there are warnings to be made about nutrition and hygiene education. It needs to be easily understood by illiterate or barely literate people. It must address issues that people can respond to. (It does little good to tell people without fuel to boil their water.) Finally, it must do more than simply inform; it must induce people to change their behavior.

We have been particularly impressed with the use of growth charts as an educational tool. In those situations where the charts were maintained simply for data keeping purposes, the reliability of the measurement deteriorated over time, and the cost in both time and money probably was not worth the benefit. On the other hand, the charts proved extremely

effective when incorporated as part of a nutrition education program, and were actively used to monitor the progress of the individual child.

We have adopted a broad definition of weight chart, one that embraces the concept of using an appropriate device, not necessarily the written word, to convey the messages inherent in a child's growth progress. In literate societies, the graphic chart with written explanations and instructions can work well. However, in societies where literacy is low, other devices (for example, color-coded flags) can be used to help the mother understand her child's progress.

In some circumstances, the weight chart is best kept by the mother in the home. We have seen several instances where the mother and family treat them as cherished possessions which "tell the story of their child." In this case, health-related messages included on the card are easily reinforced through education programs and disseminated through the mother to her family. In other settings, the weight chart might best be stored at a health post or center, and used as the triggering device for discussion and learning during consultation.

POLICY IMPLICATIONS

Our comments on the substantive elements of nutrition intervention were made, in part, to illustrate the importance of context in defining strategy. This should serve to reinforce the concept that the process of intervention described earlier is the only general prescription for success. The implications of this concept are many and, like the analytic results derived from quantitative data on nutrition, often subject to multiple interpretations.

One implication of the importance of process in community-level intervention is that nation-wide solutions to nutrition problems do not and cannot exist. To accommodate the varied personalities of communities or, perhaps, regions, individually tailored programs are needed. This is not meant to imply that national nutrition is irrelevant. Rather, it does suggest the need for national planners to consider local character in evolving multiple strategies for combatting malnutrition.

A second implication is that national surveillance systems are hardly likely to identify the true causes and effects of existing nutrition programs. If local input is needed to interpret detailed data from a community, a national system that gathers data at some central location is clearly going to overlook the complex sets of circumstances producing local change. This is not to say that national surveillance systems are a waste of effort. Rather it suggests the need to pattern national systems as collections of local systems tied to the diverse interventions in place at the local level. Data should be made available and used as part of local or regional interventions to help shape those interventions. Locally generated data bases originating in conjunction with particular interventions should serve as the building blocks of national surveillance systems.

Finally, a third implication of the need to shape nutrition programs in the context of local conditions is that funding procedure for interventions may need modification. The current strategy employed by many funding organizations to ensure accountability of their funds is to encourage potential intervenors to map out all the characteristics and costs of the intervention prior to initiation. Under these circumstances, modifying an intervention in response to lessons learned about the local environment becomes cumbersome at best. This is not meant to imply that funding agents need give up control over the expenditures made with their money. It does suggest the need for a better review process during an intervention--a process that makes continuing funding for any given program contingent on demonstration of additional need in view of the lessons learned.

Thus, we have come full circle: the mechanism to continuously identify those factors contributing to a program's success, also allows for accountability of program funding. Furthermore, the process by which intervention results are accurately obtained is precisely the same process which enhances program success.

BIBLIOGRAPHY

Anderson, Mary Ann. CARE preschool nutrition project: Phase II report. New York: CARE, 1977.

Austin, James E. et al. Nutrition intervention assessment and guidelines. Cambridge, Massachusetts: Harvard Institute for International Development, 1978.

Beaton, G.H. and Ghassemi, H. Supplementary feeding programmes for young children in developing countries. Report prepared for UNICEF and ACC Sub-committee on nutrition of the United Nations. October, 1979.

Chandra, R.K.; Newberne, P.M. Nutrition, immunity, and infection: Mechanisms of interactions. New York: Plenum Press, 1977.

Community Systems Foundation. An evaluation report of the P.L. 480 title II program in India. Ann Arbor, Michigan: Community Systems Foundation, 1978.

__________. Survey methods in developing countries: A planning and evaluation tool. A report to Save the Children on the Dominican Republic Survey. Ann Arbor, Michigan: Community Systems Foundation, 1978.

Drake, William D.; Fajardo, Luis F. The Promotora Program in Candelaria: A Colombian attempt to control malnutrition and disease, 1968-74. Ann Arbor, Michigan: Community Systems Foundation, 1976.

Fomon, Samuel J. Nutrition disorders of children: Prevention, screening, and followup. Rockville, Maryland: Public Health Service, U.S. Department of Health, Education, and Welfare, 1977.

Gershoff, Stanley N. et al. "Amino acid fortification of rice studies in Thailand. I. Background and baseline data." American Journal of Clinical Nutrition 28(2):February, 1975.

__________. "Nutrition studies in Thailand. II. Effects of fortification of rice with lysine, threonine, thiamin, riboflavin, vitamin A, and iron on preschool children." American Journal of Clinical Nutrition 30(7):July, 1977.

Gomez, F.; Galvan, C.R.; Cravioto, J.; Frenk, S. "Malnutrition in infancy and childhood, with special reference ot kwashiorkor." Advances in Pediatrics 7, 1955.

Habicht, J.P. and Butz, W.P. "Measurement of health and nutrition effects of large-scale nutrition intervention projects." Evaluating the Impact of Nutrition and Health Programs. Edited by Robert E. Klein. Plenum Press: New York, 1979.

Hamill, Peter V.V.; Drizd, Terence A; Johnson, Clifford L.; Reed, Robert B.; and Roche, Alex F. NCHS growth charts, 1976. Monthly vital statistics report, Vol. 25, No. 3 Supplement, June 22, 1976. National Center for Health Statistics, U.S. Department of Health, Education and Welfare.

Heller, Peter S.; Drake, William D. "Malnutrition, child morbidity, and the family decision process." Journal of Developmental Economics. North Holland Publishing Company: Amsterdam, Netherlands, 1979 (in press).

Ifekwunigwe, Aaron E. "Socio-Economic factors in the nutrition of the rural poor." Paper presented at the XI International Congress of Nutrition at Rio de Janeiro, Brazil, 1978.

Isenman, Paul J.; Singer, H.W. "Food Aid: Disincentive effects and their policy implications." Economic Development and Cultural Change 25(2):January, 1977.

Jelliffe, Derrick B. "Cultural aspects of nutrition." Practice of Pediatrics. Edited by Vincent Kelley. Hagerstown, Maryland: Harper and Row, 1975.

__________. Infant nutrition in the subtropic and tropics. WHO Monograph No. 25. Geneva: World Health Organization, 1955.

__________. The assessment of the nutritional status of the community. WHO Monograph No. 53. Geneva: World Health Organization, 1966.

Jones, Edna M.; Munger, Sara J.; Altman, James W. A field guide for evaluation of nutrition education. An experimental approach to determination of effects on food behavior in lesser developed countries. Allison Park, Pennsylvania: Synectics Corporation, 1975.

Karlin, Barry. The state of the art of delivering low cost health services in less developed countries: A summary study of 180 health projects. Washington, D.C.: American Public Health Association, 1977.

Keller, W.; Donoso, G.; DeMaeyer, E.M. "Anthropometry in nutritional surveillance: A review based on results of the WHO collaborative study on nutritional anthropometry." Nutrition Abstracts and Reviews 46(8):1976.

Lechtig, Aaron, et al. "Maternal nutrition and fetal growth in developing societies." American Journal of Diseases of Children 129(4):April, 1978.

Martorell, R.; Yarbrough, C.; Klein, R.E.; and Lechtig, A. "Malnutrition,

body size, and skeletal maturation: Interrelationships and implications for catch-up growth." Human Biology 51(3):September, 1979.

Moghissi, Kamran S.; Evans, Tommy N. Nutritional impacts in women throughout life with emphasis on reproduction. Hagerstown, Maryland: Harper and Row, 1977.

Morss, Elliott R.; Hatch, John K.; Mickelwait, Donald R.; Sweet, Charles F. Strategies for small farmer development. Colorado: Westview Press, Inc., 1976.

Mosely, W. Henry, ed. Nutrition and human reproduction. New York: Plenum Press, 1978.

Nelson, W., ed. Textbook of pediatrics. 9th ed. Philadelphia: W.B. Saunders, 1969.

Pollitt, Ernesto; Gilmore, Michael; Valcarcel, Martha. "Maternal and infant behaviors regulating early growth." Nutrition in transition: Proceedings, western hemisphere nutrition Congress V. Chicago: American Medical Association, 1978.

Robson, John R.K. Malnutrition: Its causation and control. New York: Gordon and Breach, 1972.

Rubenstein, A. et al. "Effect of improved sanitary facilities on infant diarrhea in a Hopi village." Public Health Reports 84(13):December, 1969.

Scrimshaw, N.S.; Taylor, C.E.; Gordon, J.E. Interactions of nutrition and infection. Geneva: World Health Organization, 1968.

Suskind, Robert M., ed. Malnutrition and the immune response. Kroc Foundation Series. Vol. 7. New York: Raven Press, 1977.

Urrutia, Juan Jose. Corn fortification: A field demonstration model. Guatemala City: Institute of Nutrition of Central America and Panama (INCAP), 1976.

Valverde, Victor et al. "Relationship between family land availability and nutritional status." Ecology of Food and Nutrition 6(1):January, 1977.

Waterlow, J.C. "Classification and definition of protein-calorie malnutrition." Medical Practice 3:September, 1972.

Zeitlin, Marian Frank. Breast-feeding and growth status in two preschool child populations; reference standards for maximum thigh circumference from 0-5 years; thigh circumference - An age-independent screening index for preschool nutrition interventions. Ph.D. dissertation, Massachusetts Institute of Technology, 1977.

[illegible]

[illegible], Carolyn S.; [illegible], Thomas M. [illegible] with emphasis on reproduction. [illegible]

[illegible] Strategies for small farmer development. [illegible]

[illegible] Nutrition and human reproduction. [illegible] Plenum Press, [illegible]

Nelson, W. ed. Textbook of pediatrics. 9th ed. Philadelphia, W.B. Saunders, 1969.

Pollitt, Ernesto; Gilmore, Michael; Valcarcel, Marcela. "Maternal [illegible] infant [illegible] growth", Nutrition in transition. Proceedings Western Hemisphere Nutrition Congress V. [illegible]

Robson, John R.K. Malnutrition: its causation and control. New York: Gordon and Breach, 1972.

Rubenstein, A. et al. "Effect of improved sanitary facilities on infant diarrhea in a Hopi village." Public Health Reports 84(12):December, 1969.

Scrimshaw, N.S.; Taylor, C.E.; Gordon, J.E. Interactions of nutrition and infection. Geneva: World Health Organization, 1968.

Suskind, Robert M. ed. Malnutrition and the immune response. Kroc Foundation Series. Vol. 7. New York: Raven Press, 1977.

[illegible], 1976.

[illegible]

Waterlow, J.C. "Classification and definition of protein-calorie malnutrition." British Medical Journal, [illegible]

[illegible] reference standards for maximum [illegible]